

Springer Proceedings in Physics 273

Aida Binti Mustapha
Suhadir Shamsuddin
Syed Zuhaib Haider Rizvi
Saliza Binti Asman
Siti Suhana Jamaian *Editors*

Proceedings of the 7th International Conference on the Applications of Science and Mathematics 2021

Sciematic 2021

 Springer

Springer Proceedings in Physics

Volume 273

Indexed by Scopus

The series Springer Proceedings in Physics, founded in 1984, is devoted to timely reports of state-of-the-art developments in physics and related sciences. Typically based on material presented at conferences, workshops and similar scientific meetings, volumes published in this series will constitute a comprehensive up-to-date source of reference on a field or subfield of relevance in contemporary physics. Proposals must include the following:

- name, place and date of the scientific meeting
- a link to the committees (local organization, international advisors etc.)
- scientific description of the meeting
- list of invited/plenary speakers
- an estimate of the planned proceedings book parameters (number of pages/articles, requested number of bulk copies, submission deadline).

Please contact:

For Americas and Europe: Dr. Zachary Evenson; zachary.evenson@springer.com
For Asia, Australia and New Zealand: Dr. Loyola DSilva; loyola.dsilva@springer.com

More information about this series at <https://link.springer.com/bookseries/361>

Aida Binti Mustapha · Suhadir Shamsuddin ·
Syed Zuhaib Haider Rizvi · Saliza Binti Asman ·
Siti Suhana Jamaian
Editors

Proceedings of the 7th International Conference on the Applications of Science and Mathematics 2021

Sciematic 2021

 Springer

Editors

Aida Binti Mustapha
Universiti Tun Hussein Onn Malaysia
Johor, Malaysia

Suhadir Shamsuddin
Universiti Tun Hussein Onn Malaysia
Johor, Malaysia

Syed Zuhaib Haider Rizvi
Universiti Tun Hussein Onn Malaysia
Johor, Malaysia

Saliza Binti Asman
Universiti Tun Hussein Onn Malaysia
Johor, Malaysia

Siti Suhana Jamaian
Universiti Tun Hussein Onn Malaysia
Johor, Malaysia

ISSN 0930-8989

ISSN 1867-4941 (electronic)

Springer Proceedings in Physics

ISBN 978-981-16-8902-4

ISBN 978-981-16-8903-1 (eBook)

<https://doi.org/10.1007/978-981-16-8903-1>

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2022

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd. The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Conference Organisation

Patron

Wahid Razzaly, Vice-Chancellor Universiti Tun Hussein Onn Malaysia

Adviser

Hashim Saim, Universiti Tun Hussein Onn Malaysia

Chairperson

Zalilah Murni Yunus, Universiti Tun Hussein Onn Malaysia

Deputy Chairman

Fahmiruddin bin Esa, Universiti Tun Hussein Onn Malaysia

Secretary

Nurun Najwa Ruslan, Universiti Tun Hussein Onn Malaysia
Noor Azura Awang, Universiti Tun Hussein Onn Malaysia

Treasurer

Nurul Nadia Adnan, Universiti Tun Hussein Onn Malaysia
Nur Amirah Liyana bijti Mohd Lazim, Universiti Tun Hussein Onn Malaysia

Publication

Syed Zuhaib Haider Rizvi, Universiti Tun Hussein Onn Malaysia
Muhammad Sufi bin Roslan, Universiti Tun Hussein Onn Malaysia
Ahmad Hassan Sallehudin bin Mohd Sarif, Universiti Tun Hussein Onn Malaysia

Manuscript Review & Editorial

Ahmad Hadi Ali, Universiti Tun Hussein Onn Malaysia
Suhadir Shamsuddin, Universiti Tun Hussein Onn Malaysia
Syed Zuhaib Haider Rizvi, Universiti Tun Hussein Onn Malaysia
Saliza Binti Asman, Universiti Tun Hussein Onn Malaysia
Siti Suhana Jamaian, Universiti Tun Hussein Onn Malaysia
Khuneswari A/P Gopal Pillay, Universiti Tun Hussein Onn Malaysia
Mohammad Ahyar bin Zulkefli, Universiti Tun Hussein Onn Malaysia

Protocol & Inauguration

Hatijah Basri, Universiti Tun Hussein Onn Malaysia
Mohamad Khidir b Mohd Ibrahim, Universiti Tun Hussein Onn Malaysia
Norzieyana bt Md Arshad, Universiti Tun Hussein Onn Malaysia
Shuhaida Binti Ismail, Universiti Tun Hussein Onn Malaysia

Promotion

Amira Saryati Ameruddin, Universiti Tun Hussein Onn Malaysia
Faridah Abu Bakar, Universiti Tun Hussein Onn Malaysia
Afishah Alias, Universiti Tun Hussein Onn Malaysia
Mohd Asrul Affendi bin Abdullah, Universiti Tun Hussein Onn Malaysia

Sponsorship, Collaboration & Keynote Speaker

Shakila Abdullah, Universiti Tun Hussein Onn Malaysia
Nur Hafizah binti Malik, Universiti Tun Hussein Onn Malaysia
Mohd Fadzelly bin Abu Bakar, Universiti Tun Hussein Onn Malaysia

Registration & Certificates

Siti Amira Othman, Universiti Tun Hussein Onn Malaysia
Cik Issha Nadirah Ismail, Universiti Tun Hussein Onn Malaysia
Cik Siti Maisarah Rahim, Universiti Tun Hussein Onn Malaysia

Program book & Souvenir

Syazwan Hanani Meriam Suhaimy, Universiti Tun Hussein Onn Malaysia
Faridah Kormin, Universiti Tun Hussein Onn Malaysia
Munira Zainal Abidin, Universiti Tun Hussein Onn Malaysia

Technical & Logistics

Mohd Zul Hilmi bin Mayzan, Universiti Tun Hussein Onn Malaysia
Mohd Marhafidz Marjori, Universiti Tun Hussein Onn Malaysia
Nooriskandar Sani, Universiti Tun Hussein Onn Malaysia
Tajul Asmawee Abdullah, Universiti Tun Hussein Onn Malaysia
Sufian Abd. Rahim, Universiti Tun Hussein Onn Malaysia
Kamarul Affendi Hamdan, Universiti Tun Hussein Onn Malaysia
Hemmy Abd Jalal, Global, Universiti Tun Hussein Onn Malaysia
Mustaffa Kamal Md Saim, PTM, Universiti Tun Hussein onn Malaysia

Technical Sessions

Suzi Salwah Jikan, Universiti Tun Hussein Onn Malaysia
Norhaidah Binti Mohd Asrah, Universiti Tun Hussein Onn Malaysia
Siti Fatimah Zaharah Binti Mohamad Fuzi, Universiti Tun Hussein Onn Malaysia
Muhamad Muizzudin Azali, Universiti Tun Hussein Onn Malaysia

Editorial Board

Chief Editor

Ahmad Hadi Ali, Universiti Tun Hussein Onn Malaysia

Advisory Editorial Board

Ishak Hashim, Universiti Kebangsaan Malaysia
Sharifah Mohamad, Universiti Malaya
Hussein Yousif Eledum, Tabuk University, Saudi Arabia
Tom G. Mackay, University of Edinburgh, United Kingdom
Hosam Hegazy, Jazan University, Saudi Arabia
Darmesah Gabda, Universiti Malaysia Sabah

Editorial Board

Siti Suhana Jamaian, Universiti Tun Hussein Onn Malaysia
Saliza Binti Asman, Universiti Tun Hussein Onn Malaysia
Suhadir Shamsuddin, Universiti Tun Hussein Onn Malaysia
Khuneswari Gopal Pillay, Universiti Tun Hussein Onn Malaysia
Syed Zuhaib Haider Rizvi, Universiti Tun Hussein Onn Malaysia
Abd. El-Raheem Mohamed, Ain Shams University, Egypt
Randa Alharbi, Tabuk University, Saudi Arabia

Proceedings Editors

Aida Binti Mustapha, Universiti Tun Hussein Onn Malaysia
Siti Suhana Jamaian, Universiti Tun Hussein Onn Malaysia
Saliza Binti Asman, Universiti Tun Hussein Onn Malaysia
Suhadir Shamsuddin, Universiti Tun Hussein Onn Malaysia
Syed Zuhaib Haider Rizvi, Universiti Tun Hussein Onn Malaysia
Zalilah Murni Yunus, Universiti Tun Hussein Onn Malaysia
Hazel Monica Matias-Peralta, Central Luzon State University, Philippines

Preface

From the smallest of atoms to the largest of real numbers, science and mathematics are sorely needed in solving many of the major contemporary issues: energy, climate change, poverty, big data, food and water. The world is looking for sustainable solutions to problems more than ever. The synergistic approach of Mathematicians, Scientists and Engineers has undeniable importance for finding and implementing sustainable solutions to our biggest problems. Nevertheless, as changes occur and development takes place, the responsibility of sustaining it does not only lie in the hands of scientists, mathematicians and engineers but also on the society. People at large, irrespective of their backgrounds, play the most crucial roles in ensuring that the world we live in would continue to flourish. Hence, it is imperative for us to disseminate facts to the public with thorough honesty. They should be enlightened with the latest developments in the field of science and engineering which are aimed for sustainable future technologies.

With this viewpoint SCIEMATHIC 2021 is planned with a theme “Quest for Sustainable Science and Mathematics for Future Technologies.” It is the seventh in the series of SCIEMATHIC conferences organised by Faculty of Applied Sciences and Technology (FAST), Universiti Tun Hussein Onn Malaysia every year. As COVID-19 precautionary measure, it is organised online this year during 27–28 October 2021.

This book highlights the work of amazing researchers from different parts of the world with a focus on potential applications for future technologies. It will provide a quality insight to the reader about the leading trends in sustainable science and technology research. It covers the research work of Physicists, Mathematicians, Statisticians and Data Scientists, aiming to find sustainable solutions to the major problems in the scientific world.

From SCIEMATHIC 2021, we attracted papers from 6 different countries. In this volume of SPP we are including 39 high quality papers, each of these is accepted for oral presentation and publication after blind peer review by at least two independent reviewers. The papers in this proceeding are divided into four sections Physics, Mathematics, Statistics and Engineering, as they were presented during the conference.

On behalf of the SCIEATHIC 2021, we express our highest gratitude to the Patron Prof. Datuk Ts. Dr. Wahid Razzaly; Advisor, Prof. Dr. Hashim Saim; Conference Organiser, Faculty of Applied Sciences & Technology, UTHM, Conference Chair, Publication Committee, Manuscript & Review Committee, and all Committee members for their valuable efforts in the review process that helped us to guarantee the highest quality of the selected papers for the conference.

We would like to express our thanks to the Keynote speakers: Dato' Sri Ibrahim Ahmad from Universiti Tun Hussein Onn Malaysia, Asst. Prof. Dr. Satoru Yoshioka from Kyushu University, Japan, Dr. John Pasley from University of York, England, Dr. Muhammad Shahbaz Anwar from University of Cambridge, England, and ChM. Dr. Siti Fatimah Salim from Universiti Teknologi Mara, Malaysia.

Invited Speakers: Prof. Dr. Shafida Abd Hamid, International Islamic University Malaysia, Dr. Saiful Najmee bin Mohammad, Universiti Teknologi Mara and Mr. Abdul Rahim bin Othman, PETRONAS for sharing their inspiring talks.

We would also like to say thanks to all the committee members and session chairs for their substantial work in making this event successful.

Our special thanks to Dr. Loyol D'Silva and Ms. Shalini Monica for publishing this proceeding in Springer Proceedings in Physics.

Lastly, we cordially thank all the authors for their valuable contributions and other participants of this conference. The conference would not have been possible without them.

Johor, Malaysia

Aida Binti Mustapha
Suhadir Shamsuddin
Syed Zuhaib Haider Rizvi
Saliza Binti Asman
Siti Suhana Jamaian

Contents

Part I Physics

1	Passively Q-Switched Erbium-Doped Fiber Laser Based on PSF-ZnO-TiO₂ Nanoparticles as Saturable Absorber	3
	Nik N. H. E. N. Mahmud, Noor A. Awang, Nurul I. Rasli, and Hatijah Basri	
2	Validation and Thickness Dependency of a Planar Material Return Loss by Numerical Simulation	17
	Wan Azmafaliyana Wan Azmi and Fahmiruddin Esa	
3	Analytical Simulation of Spatial Distribution Lambertian Point-Source in LED Chip's Epitaxial Layers	27
	Ahmad Noor Abdul Hamid, Ahmad Fakhurrazi Ahmad Noorden, and Muhammad Zamzuri Abdul Kadir	
4	Free Spectral Range, and Coupling Coefficient Analysis of Integrated Fano Resonance in All-Pass Micro-resonator	35
	Mohammad Amirul Hairol Aman, Ahmad Fakhurrazi Ahmad Noorden, Faris Azim Ahmad Fajri, Ahmad Noor Abdul Hamid, and Suzairi Daud	
5	Doping Effect on Bandgap Energy and Luminescence Spectrum for AlN-Based Semiconductor	43
	Faris Azim Ahmad Fajri, Mohammad Amirul Hairol Aman, Ahmad Fakhurrazi Ahmad Noorden, Ahmad Noor Abdul Hamid, and Azni Abdul Aziz	
6	Characterization of Electrical, Optical and Topological Properties of ITO Thin Films for Solar Cells Application	57
	Nursyahrah Masdan and Ahmad Hadi Ali	

7	The Effect of pH on Particle Size of Hydrothermally Reduced Graphene Oxide	65
	Yi Lin Chan, Fahmiruddin Esa, Kok Yeow You, and Man Seng Sim	
8	A Comparative Study of Cr and Mo Ultrathin Films as Transparent Conductors in Solar Cell Application	73
	Abdelbaki Hacini, Ahmad Hadi Ali, and Nurul Nadia Adnan	
9	Simulation on Optical Absorption for Amorphous Silicon Thin Film Solar Cell with CdSe/ZnS Quantum Dots	81
	Mirza Basyir Rodhuan, Rosmila Abdul-Kahar, and Amira Saryati Ameruddin	
10	Electronic Structure and Optical Properties of Lead-Free Double Perovskite BaMgLaBiO₆: A DFT Study	95
	Junaid Munir, Muhammad Jamil, Quratul Ain, and Kaneez Fatima	
11	Conductivity of PANI/ZnO Nanocomposites	107
	Siti Ashikin Hassan, Maytham Qabel Hamzah, and Mohd Arif Agam	
12	A Mach-Zehnder Ring Resonator Layout for Label Free Detection of Diabetes	115
	Parisa Sanati, Mahdi Bahadoran, and Mohd Kamarulzaki Mustafa	
13	Natural Radioactivity, Transfer Factor and Associated Radiological Risk in Commercially Cultivated Yam (<i>Dioscorea Rotundata</i>) in Northcentral Nigeria	125
	Matthew Tikpangi Kolo, Oyeleke Ismail Olarinoye, Simon Olonkwoh Saliyu, Hauwau Kulu Shuaibu, and Funmilayo Ayedun	
14	Analysis of Electrical Transport Properties in Nd_{0.75}Na_{0.25-x}K_xMnO₃ Manganites	139
	S. Shamsuddin, N. Ibrahim, M. Z. H. Mayzan, Fahmiruddin Esa, and S. A. Razali	
15	Crystalline Structure, Surface Morphology of Perovskite Manganites Nd_{0.75}Na_{0.25}Mn_{1-y}Fe_yO₃	147
	S. Shamsuddin, M. Z. H. Mayzan, Fahmiruddin Esa, and L. K. Hao	
16	Element Identification of Different Gemstones by Using LIBS-PCA Method	155
	Noradilah Mohamad Saufi, Nurul Nadia Adnan, Syed Zuhaib Haider Rizvi, Mohd Sufi Roslan, Siti Norazida Razak, Nurhafizah Hasim, and Nurin Iffah Mansor	

17 Discrimination of Precious and Semi-Precious Gemstones Using Laser-Induced Breakdown Spectroscopy and Machine Learning Approaches 165
 Ain Zamira, Syed Zuhaib Haider Rizvi, Mohd Sufi Roslan, Nurul Nadia Adnan, Siti Noraiza Ab Razak, and Muhammad Arif Jalil

18 Structural Properties of Lithium-Niobium-Calcium Tellurite Glass Doped with Nickel Oxide: Computational Analysis 175
 S. K. Farisya Nadira, H. Nurhafizah, N. H. Ahmad, and Nurul Nadia Adnan

Part II Mathematics

19 Comparing the Hamming and Manhattan Heuristics in Solving the 8—Puzzle by A* Algorithm 189
 Siaw Chong Lee and Tyan Her See

20 Magneto hydrodynamic Effects in Mixed Convection Viscoelastic Fluid Flow at Lower Stagnation Point on a Sliced Sphere 197
 Basuki Widodo, Tunjung Eviloka, Chairul Imron, and Dieky Adzkiya

21 Stability Analysis of Fractional-Order Chemostat Model with Time Delay 213
 Nor Afiqah Mohd Aris and Siti Suhana Jamaian

22 Free Convection Nanofluid Flow Near a Three-Dimensional Stagnation Point Induced by g-Jitter: Constant Heat Flux 229
 Mohamad Hidayad Ahmad Kamal, Anati Ali, Noraihan Afiqah Rawi, and Sharidan Shafie

23 Spatial Decision Support System (SDSS) for Sustainable Business Location Selection of Purpose-Built Offices (PBO) in Malaysia 241
 Jia Yun Heng, Edie Ezwan Mohd Safian, and Burhaida Burhan

Part III Statistics

24 Investigation on the Effect of Household Income and Demographic Factors on Healthcare Cost in Perak 251
 Khuneswari Gopal Pillay and Wong Pui Yee

25 Classification Tree of Breast Cancer Data with Mode Value for Missing Data Replacement 265
 Nur Atiqah Hamzah, Sabariah Saharan, and Khuneswari Gopal Pillay

26 Socio Demographic Determinants of Contraceptive Use Among Women in Malaysia: Logistic Regression Model 279
 Nor Azah Samat, Noor Azlin Muhammad Sapri,
 Norziha Che Him, Adzmel Mahmud, and Noor Wahida Md. Junus

27 Crude Palm Oil Price Forecasting: Comparative Study of Hybrid GMDH Neural Network and ARIMA Model 289
 Huma Basheer and Azme Khamis

28 Analyzing of Traffic Accidents in Skudai City Using Stochastic Models 309
 Teng Mun Jing and Rohayu Mohd Salleh

29 The FGVH and GDEX Stock Price Modeling Using ARIMA and Holt’s Linear Trend Methods 325
 Norhaidah Mohd Asrah, Rabia’tul Adawiyah Sua’aif,
 and Nur Hazlina Abdul Wahab

30 Deep Learning Approach for Football Match Classification of English Premier League (EPL) Based on Full-Time Results 339
 Muhaimin Muszaidi, Aida Binti Mustapha, Shuhaida Ismail,
 and Nazim Razali

31 Comparative Analysis of Statistical and Machine Learning Methods for Classification of Match Outcomes in Association Football 351
 Syazira Zulkifli, Aida Binti Mustapha, Shuhaida Ismail,
 and Nazim Razali

Part IV Engineering

32 Pre-assessment of Machine Learning Approaches for Patient Length of Stay Prediction 369
 Farhanahani Mahmud, Ahmad Zahran Md Khudzari,
 Cheong Ping Pau, Mohd Faizal Ramli, Norfazlina Jaffar,
 and Intan Fariza Gaaffar

33 Prediction Model of Mass Rapid Transit Noise Level Using the Stepwise Regression Analysis 379
 Nor Baizura Hamid, Mohd Erwan Sanik,
 Hafsa Mohammad Noor, Joewono Prasetijo, Mardiha Mokhtar,
 Mohamad Azim Mohammad Azmi, Mohamad Irwan Yahaya,
 and Mohd Zakwan Ramli

34 Engine Monitoring During Tuning Process for Diesel-Compressed Natural Gas (CNG) Dual Fuel Engine Using Statistical Approach 391
 Abd Fathul Hakim Zulkifli, Mas Fawzi Mohd Ali,
 Muammar Mukhsin Ismail, and Shahrul Azmir Osman

35 Impulse Breakdown Characteristics of CF₃I–CO₂ at Various Gas Pressure and CF₃I Content 405
Muhammad Saufi Kamarudin, Noor Mazliza Badrul Sham,
A. Haddad, Md Nor Ramdon Baharom,
Mohd Fairouz Mohd Yousof, and Nordiana Azlin Othman

36 Optimal Rain Attenuation Prediction Models for Earth-Space Communication at Ku-Band in North Central Nigeria 415
K. C. Igwe

37 Investigation on Water Losses in Reticulation System at UTHM Campus 429
Sumitha Selami and Sabariah Musa

38 Short-Term Time-Series Forecasting by Using Exponential Smoothing Techniques at Sungai Pahang, Malaysia 443
Ahmad Murshid Hamidon and Sabariah Musa

39 Internet of Things for Hybrid Energy System Data Monitoring 453
Saadatul Suhaili, Faridahanim Ahmad, Tsu Yian Lee,
Nur Izieadiana Abidin, and Mariah Awang

Contributors

Azni Abdul Aziz Centre for Advanced Optoelectronics Research (CAPTOR), Kulliyah of Science, International Islamic University Malaysia, Pahang, Malaysia

Ahmad Noor Abdul Hamid Centre for Advanced Optoelectronics Research (CAPTOR), Kulliyah of Science, International Islamic University Malaysia, Pahang, Malaysia

Muhammad Zamzuri Abdul Kadir Centre for Advanced Optoelectronics Research (CAPTOR), Kulliyah of Science, International Islamic University Malaysia, Pahang, Malaysia

Rosmila Abdul-Kahar Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, Educational Hub Pagoh, Muar, Johor, Malaysia

Nur Izieadiana Abidin Department of Structure and Material, School of Civil Engineering, Faculty of Engineering, Universiti Teknologi Malaysia, Bahru, Johor, Malaysia

Nurul Nadia Adnan Laser and Semiconductor Technology Research Group, Center of Research (CoR) Photonics Devices and Sensors Research Center (PDSR), Department of Physics and Chemistry, Faculty of Applied Sciences and Technology, Pagoh Educational Hub, Universiti Tun Hussein Onn Malaysia, Pagoh, Johor, Malaysia

Dieky Adzkiya Department of Mathematics, Institut Teknologi Sepuluh Nopember, Kampus ITS Sukolilo, Surabaya, Indonesia

Mohd Arif Agam Department of Physics, Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, Pagoh, Johor, Malaysia

Faris Azim Ahmad Fajri Centre for Advanced Optoelectronics Research (CAPTOR), Kulliyah of Science, International Islamic University Malaysia, Pahang, Malaysia

Ahmad Fakhurrizi Ahmad Noorden Centre for Advanced Optoelectronics Research (CAPTOR), Kulliyah of Science, International Islamic University Malaysia, Pahang, Malaysia

Faridahanim Ahmad Department of Structure and Material, School of Civil Engineering, Faculty of Engineering, Universiti Teknologi Malaysia, Bahru, Johor, Malaysia

N. H. Ahmad Advanced Optical Material Research Group (AOMRG), Department of Physics, Faculty of Science, Universiti Teknologi Malaysia, Bahru, Johor, Malaysia

Quratul Ain Department of Physics, University of Management and Technology, Lahore, Pakistan

Ahmad Hadi Ali Laser and Semiconductor Technology Research Group (LASERG), PDRS-COR, Department of Physics and Chemistry, Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, Johor, Malaysia

Anati Ali Department of Mathematical Sciences, Faculty of Science, Universiti Teknologi Malaysia (UTM), Johor Bahru, Johor Darul Takzim, Malaysia

Mas Fawzi Mohd Ali Automotive Research Research Group, Centre for Energy and Industrial Environment Studies (CEIES), Faculty of Mechanical and Manufacturing Engineering, Universiti Tun Hussein Onn Malaysia, Batu Pahat, Johor, Malaysia

Amira Saryati Ameruddin Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, Educational Hub Pagoh, Muar, Johor, Malaysia

Nor Afiqah Mohd Aris Department of Mathematics and Statistics, Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, Pagoh, Johor, Malaysia

Norhaidah Mohd Asrah Data Analytics, Sciences and Modeling Focus Group, Universiti Tun Hussein Onn Malaysia, Pagoh, Johor, Malaysia;
Department of Mathematics and Statistics, Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, Pagoh, Johor, Malaysia

Mariah Awang Department of Civil Engineering Technology, Faculty of Engineering Technology, Universiti Tun Hussein Onn Malaysia (Pagoh Campus), Batu Pahat, Johor, Malaysia

Noor A. Awang Optical Fiber Laser Technology (OpFLAT) Focus Group, Department of Physics and Chemistry, Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, Pagoh, Johor, Malaysia

Funmilayo Ayedun Department of Physics, National Open University of Nigeria, Abuja, Nigeria

Mohamad Azim Mohammad Azmi Spatial Technology for Civil Engineering (STforCE), Centre for Diploma Studies (CeDS), Universiti Tun Hussein Onn Malaysia, Pagoh Education Hub, Pagoh, Johor, Malaysia

Wan Azmafaliyana Wan Azmi Department of Physics and Chemistry, Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, Pagoh, Muar, Johor, Malaysia

Noor Mazliza Badrul Sham Department of Electrical Power, Faculty of Electrical and Electronic Engineering, Universiti Tun Hussein Onn Malaysia (UTHM), Batu Pahat, Johor, Malaysia

Mahdi Bahadoran Department of Physics, Shiraz University of Technology, Shiraz, Fars, Iran

Md Nor Ramdon Baharom Department of Electrical Power, Faculty of Electrical and Electronic Engineering, Universiti Tun Hussein Onn Malaysia (UTHM), Batu Pahat, Johor, Malaysia

Huma Basheer Department of Mathematics and Statistics, Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia (UTHM) Edu Hub Pagoh, Pagoh, Johor, Malaysia

Hatijah Basri Optical Fiber Laser Technology (OpFLAT) Focus Group, Department of Physics and Chemistry, Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, Pagoh, Johor, Malaysia

Burhaida Burhan Department of Real Estate Management, Faculty of Technology & Business Management, Universiti Tun Hussein Onn Malaysia, Parit Raja, Johor, Malaysia

Yi Lin Chan Department of Physics and Chemistry, Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, Pagoh, Muar, Johor, Malaysia

Suzairi Daud Laser Center, Ibnu Sina Institute for Scientific and Industrial Research, Universiti Teknologi Malaysia, Johor, Malaysia

Fahmiruddin Esa Department of Physics and Chemistry, Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, Pagoh, Muar, Johor, Malaysia;
Ceramic and Amorphous Group, Faculty of Applied Sciences and Technology, Pagoh Higher Education Hub, Universiti Tun Hussein Onn Malaysia, Panchor, Johor, Malaysia

Tunjung Eviloka Department of Mathematics, Institut Teknologi Sepuluh Nopember, Kampus ITS Sukolilo, Surabaya, Indonesia

S. K. Farisya Nadira Advanced Optical Material Research Group (AOMRG), Department of Physics, Faculty of Science, Universiti Teknologi Malaysia, Bahru, Johor, Malaysia

Kaneez Fatima Department of Physics, Riphah International University, Lahore, Pakistan

Intan Fariza Gaaffar Clinical Research Department (CRD), Institut Jantung Negara (IJN), Kuala Lumpur, Wilayah Persekutuan Kuala Lumpur, Malaysia

Abdelbaki Hacini Laser and Semiconductor Technology Research Group, COR PDSR, Department of Physics and Chemistry, Faculty of Applied Sciences and Technology, Pagoh Educational Hub, Universiti Tun Hussein Onn Malaysia, Pagoh, Johor, Malaysia

A. Haddad School of Engineering, Cardiff University, Cardiff, UK

Mohammad Amirul Hairol Aman Centre for Advanced Optoelectronics Research (CAPTOR), Kulliyah of Science, International Islamic University Malaysia, Pahang, Malaysia

Nor Baizura Hamid Spatial Technology for Civil Engineering (STforCE), Centre for Diploma Studies (CeDS), Universiti Tun Hussein Onn Malaysia, Pagoh Education Hub, Pagoh, Johor, Malaysia

Ahmad Murshid Hamidon Department of Civil Engineering, Faculty of Civil Engineering and Built Environment, Universiti Tun Hussein Onn Malaysia, Batu Pahat, Malaysia

Maytham Qabel Hamzah Department of Physics, Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, Pagoh, Johor, Malaysia; The General Directorate of Education, Ministry of Education, Al-Muthanna Governorate, Republic of Iraq; Department of Mathematics and Statistics, Faculty of Applied Science and Technology, Universiti Tun Hussein Onn Malaysia, Batu Pahat, Johor, Malaysia

Nur Atiqah Hamzah Department of Mathematics and Statistics, Faculty of Applied Science and Technology, Universiti Tun Hussein Onn Malaysia, Batu Pahat, Johor, Malaysia

L. K. Hao Ceramic and Amorphous Group, Faculty of Applied Sciences and Technology, Pagoh Higher Education Hub, Universiti Tun Hussein Onn Malaysia, Panchor, Johor, Malaysia

Nurhafizah Hasim Advanced Optical Material Research Group (AOMRG), Department of Physics, Faculty of Science, Universiti Teknologi Malaysia, Bahru, Johor, Malaysia

Siti Ashikin Hassan Department of Physics, Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, Pagoh, Johor, Malaysia

Jia Yun Heng Department of Real Estate Management, Faculty of Technology & Business Management, Universiti Tun Hussein Onn Malaysia, Parit Raja, Johor, Malaysia

Norzha Che Him Department of Mathematics and Statistics, Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, Pagoh, Johor, Malaysia

N. Ibrahim Faculty of Applied Sciences, Universiti Teknologi MARA, Shah Alam, Selangor, Malaysia

K. C. Igwe Department of Physics, Federal University of Technology, Niger State, Minna, Nigeria

Chairul Imron Department of Mathematics, Institut Teknologi Sepuluh Nopember, Kampus ITS Sukolilo, Surabaya, Indonesia

Muammar Mukhsin Ismail Automotive Research Research Group, Centre for Energy and Industrial Environment Studies (CEIES), Faculty of Mechanical and Manufacturing Engineering, Universiti Tun Hussein Onn Malaysia, Batu Pahat, Johor, Malaysia

Shuhaida Ismail Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, Johor, Panchor, Malaysia

Norfazlina Jaffar Clinical Research Department (CRD), Institut Jantung Negara (IJN), Kuala Lumpur, Wilayah Persekutuan Kuala Lumpur, Malaysia

Muhammad Arif Jalil Physics Department, Faculty of Science, Universiti Teknologi Malaysia, Skudai, Johor Bahru, Johor, Malaysia

Siti Suhana Jamaian Department of Mathematics and Statistics, Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, Pagoh, Johor, Malaysia

Muhammad Jamil Department of Physics, Riphah International University, Lahore, Pakistan

Teng Mun Jing Department of Mathematics and Statistics, Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, Pagoh, Johor, Malaysia

Noor Wahida Md. Junus Department of Mathematics, Faculty of Science and Mathematics, Universiti Pendidikan Sultan Idris, Tanjong Malim, Perak, Malaysia

Mohamad Hidayad Ahmad Kamal Department of Mathematical Sciences, Faculty of Science, Universiti Teknologi Malaysia (UTM), Johor Bahru, Johor Darul Takzim, Malaysia

Muhammad Saufi Kamarudin Department of Electrical Power, Faculty of Electrical and Electronic Engineering, Universiti Tun Hussein Onn Malaysia (UTHM), Batu Pahat, Johor, Malaysia

Azme Khamis Department of Mathematics and Statistics, Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia (UTHM) Edu Hub Pagoh, Pagoh, Johor, Malaysia

Ahmad Zahran Md Khudzari School of Biomedical Engineering and Health Sciences, Faculty of Engineering, IJN-UTM Cardiovascular Engineering Centre (IJN-UTM CEC), Institute of Human Centered Engineering (iHUMEN), Universiti Teknologi Malaysia (UTM), Johor Bahru, Johor, Malaysia

Matthew Tikpangi Kolo Department of Physics, Federal University of Technology, Minna, Niger State, Nigeria

Siaw Chong Lee Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, Parit Raja, Johor, Malaysia

Tsu Yian Lee Department of Structure and Material, School of Civil Engineering, Faculty of Engineering, Universiti Teknologi Malaysia, Bahru, Johor, Malaysia

Adzmel Mahmud Population and Family Research Division, National Population and Family Development Board, Kuala Lumpur, Malaysia

Farhanahani Mahmud Faculty of Electrical and Electronic Engineering, Micro-electronics and Nanotechnology Shamsuddin Research Centre (MiNT-SRC), Institute of Integrated Engineering (I2E), Universiti Tun Hussein Onn Malaysia (UTHM), Parit Raja, Batu Pahat, Johor, Malaysia

Nik N. H. E. N. Mahmud Optical Fiber Laser Technology (OpFLAT) Focus Group, Department of Physics and Chemistry, Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, Pagoh, Johor, Malaysia

Nurin Iffah Mansor Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, Pagoh Educational Hub, Pagoh, Johor, Malaysia

Nursyahirah Masdan Laser and Semiconductor Technology Research Group (LASERG), PDRS-COR, Department of Physics and Chemistry, Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, Johor, Malaysia

M. Z. H. Mayzan Ceramic and Amorphous Group, Faculty of Applied Sciences and Technology, Pagoh Higher Education Hub, Universiti Tun Hussein Onn Malaysia, Panchor, Johor, Malaysia

Edie Ezwan Mohd Safian Department of Real Estate Management, Faculty of Technology & Business Management, Universiti Tun Hussein Onn Malaysia, Parit Raja, Johor, Malaysia

Mohd Fairouz Mohd Yousof Department of Electrical Power, Faculty of Electrical and Electronic Engineering, Universiti Tun Hussein Onn Malaysia (UTHM), Batu Pahat, Johor, Malaysia

Mardiha Mokhtar Spatial Technology for Civil Engineering (STforCE), Centre for Diploma Studies (CeDS), Universiti Tun Hussein Onn Malaysia, Pagoh Education Hub, Pagoh, Johor, Malaysia

Junaid Munir Department of Physics, Riphah International University, Lahore, Pakistan

Sabariah Musa Department of Civil Engineering, Faculty of Civil Engineering and Built Environment, Eco-Hytech Research Centre, Universiti Tun Hussein Onn Malaysia, Batu Pahat, Malaysia

Mohd Kamarulzaki Mustafa Department of Physics and Chemistry, Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, Pagoh Educational Hub, Muar, Johor, Malaysia

Aida Binti Mustapha Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, Johor, Panchor, Malaysia

Muhaimin Muszaidi Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, Panchor, Johor, Malaysia

Hafsa Mohammad Noor Spatial Technology for Civil Engineering (STforCE), Centre for Diploma Studies (CeDS), Universiti Tun Hussein Onn Malaysia, Pagoh Education Hub, Pagoh, Johor, Malaysia

H. Nurhafizah Advanced Optical Material Research Group (AOMRG), Department of Physics, Faculty of Science, Universiti Teknologi Malaysia, Bahru, Johor, Malaysia

Oyeleke Ismail Olarinoye Department of Physics, Federal University of Technology, Minna, Niger State, Nigeria

Shahrul Azmir Osman Automotive Research Research Group, Centre for Energy and Industrial Environment Studies (CEIES), Faculty of Mechanical and Manufacturing Engineering, Universiti Tun Hussein Onn Malaysia, Batu Pahat, Johor, Malaysia

Nordiana Azlin Othman Department of Electrical Power, Faculty of Electrical and Electronic Engineering, Universiti Tun Hussein Onn Malaysia (UTHM), Batu Pahat, Johor, Malaysia

Cheong Ping Pau Cardiothoracic Surgery Department, Institut Jantung Negara (IJN), Kuala Lumpur, Wilayah Persekutuan Kuala Lumpur, Malaysia

Khuneswari Gopal Pillay Data Analytics, Sciences and Modeling Focus Group, Universiti Tun Hussein Onn Malaysia, Pagoh, Johor, Malaysia;
Department of Mathematics and Statistics, Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, Pagoh, Malaysia;
Department of Mathematics and Statistics, Faculty of Applied Science and Technology, Universiti Tun Hussein Onn Malaysia, Batu Pahat, Johor, Malaysia

Joewono Prasetyo Industry Centre of Excellence Railway (ICOEREL), University Tun Hussein Onn Malaysia, Pagoh Education Hub, Pagoh, Johor, Malaysia

Mohd Faizal Ramli Clinical Research Department (CRD), Institut Jantung Negara (IJN), Kuala Lumpur, Wilayah Persekutuan Kuala Lumpur, Malaysia

Mohd Zakwan Ramli College of Engineering, Universiti Tenaga Nasional, Kajang, Selangor, Malaysia

Nurul I. Rasli Optical Fiber Laser Technology (OpFLAT) Focus Group, Department of Physics and Chemistry, Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, Pagoh, Johor, Malaysia

Noraihan Afiqah Rawi Department of Mathematical Sciences, Faculty of Science, Universiti Teknologi Malaysia (UTM), Johor Bahru, Johor Darul Takzim, Malaysia

Siti Noraiza Ab Razak Center for Diploma Studies (CeDS), Universiti Tun Hussein Onn Malaysia, Pagoh, Johor, Malaysia

Siti Norazida Razak Department of Science and Mathematics, Center for Diploma Studies, Universiti Tun Hussein Onn Malaysia Educational Hub Pagoh, Muar, Johor, Malaysia

Nazim Razali Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn Malaysia, Batu Pahat, Johor, Parit Raja, Malaysia

S. A. Razali Ceramic and Amorphous Group, Faculty of Applied Sciences and Technology, Pagoh Higher Education Hub, Universiti Tun Hussein Onn Malaysia, Panchor, Johor, Malaysia

Syed Zuhair Haider Rizvi Photonics Devices and Sensor Research (PDSR), Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, Pagoh Educational Hub, Pagoh, Johor, Malaysia;

Department of Physics and Chemistry, Faculty of Applied Science and Technology (FAST), Universiti Tun Hussein Onn Malaysia, Pagoh, Johor, Malaysia;

Microelectronic and Nanotechnology Shamsuddin Research Center (MiNT-SRC), Universiti Tun Hussein Onn Malaysia, Parit Raja, Batu Pahat, Johor, Malaysia

Mirza Basyir Rodhuan Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, Educational Hub Pagoh, Muar, Johor, Malaysia

Mohd Sufi Roslan Department of Science and Mathematics, Center for Diploma Studies, Universiti Tun Hussein Onn Malaysia Educational Hub Pagoh, Muar, Johor, Malaysia;

Microelectronic and Nanotechnology Shamsuddin Research Center (MiNT-SRC), Universiti Tun Hussein Onn Malaysia, Parit Raja, Batu Pahat, Johor, Malaysia

Sabariah Saharan Department of Mathematics and Statistics, Faculty of Applied Science and Technology, Universiti Tun Hussein Onn Malaysia, Batu Pahat, Johor, Malaysia

Simon Olonkwoh Salihu Department of Chemistry, Federal University of Technology, Minna, Niger State, Nigeria

Rohayu Mohd Salleh Data Analytics, Sciences and Modeling Focus Group, Universiti Tun Hussein Onn Malaysia, Pagoh, Johor, Malaysia

Nor Azah Samat Department of Mathematics, Faculty of Science and Mathematics, Universiti Pendidikan Sultan Idris, Tanjong Malim, Perak, Malaysia

Parisa Sanati Burn and Wound Healing Research Center, Shiraz University of Medical Sciences, Shiraz, Iran

Mohd Erwan Sanik Spatial Technology for Civil Engineering (STforCE), Centre for Diploma Studies (CeDS), Universiti Tun Hussein Onn Malaysia, Pagoh Education Hub, Pagoh, Johor, Malaysia

Noor Azlin Muhammad Sapri Population and Family Research Division, National Population and Family Development Board, Kuala Lumpur, Malaysia

Noradilah Mohamad Saufi Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, Pagoh Educational Hub, Pagoh, Johor, Malaysia

Tyan Her See Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, Parit Raja, Johor, Malaysia

Sumitha Selami Department of Civil Engineering, Faculty of Civil Engineering and Built Environment, Universiti Tun Hussein Onn Malaysia, Batu Pahat, Malaysia

Sharidan Shafie Department of Mathematical Sciences, Faculty of Science, Universiti Teknologi Malaysia (UTM), Johor Bahru, Johor Darul Takzim, Malaysia

S. Shamsuddin Ceramic and Amorphous Group, Faculty of Applied Sciences and Technology, Pagoh Higher Education Hub, Universiti Tun Hussein Onn Malaysia, Panchor, Johor, Malaysia

Hauwau Kulu Shuaibu Department of Physics, Nigerian Defence Academy, Kaduna, Kaduna State, Nigeria

Man Seng Sim School of Electrical Engineering, Faculty of Engineering, Universiti Teknologi Malaysia, Skudai, Johor Bahru, Johor, Malaysia

Rabia'tul Adawiyah Sua'aif Department of Mathematics and Statistics, Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, Pagoh, Johor, Malaysia

Saadatul Suhaili Department of Structure and Material, School of Civil Engineering, Faculty of Engineering, Universiti Teknologi Malaysia, Bahru, Johor, Malaysia

Nur Hazlina Abdul Wahab Department of Mathematics and Statistics, Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, Pagoh, Johor, Malaysia

Basuki Widodo Department of Mathematics, Institut Teknologi Sepuluh Nopember, Kampus ITS Sukolilo, Surabaya, Indonesia

Mohamad Irwan Yahaya Faculty of Mechanical Engineering, Penang Branch, Universiti Teknologi Mara, Pulau Pinang, Malaysia

Wong Pui Yee Department of Mathematics and Statistics, Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, Pagoh, Malaysia

Kok Yeow You School of Electrical Engineering, Faculty of Engineering, Universiti Teknologi Malaysia, Skudai, Johor Bahru, Johor, Malaysia

Ain Zamira Center of Research (CoR) Photonics Devices and Sensors Research Center (PDSR), Universiti Tun Hussein Onn Malaysia, Pagoh, Johor, Malaysia

Abd Fathul Hakim Zulkiffi Centre of Automotive and Powertrain Technology (CAPT), Faculty of Engineering Technology, Pagoh, Muar, Johor, Malaysia

Syazira Zulkiffi Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, Johor, Panchor, Malaysia

Part I

Physics

Chapter 1

Passively Q-Switched Erbium-Doped Fiber Laser Based on PSF-ZnO-TiO₂ Nanoparticles as Saturable Absorber



Nik N. H. E. N. Mahmud, Noor A. Awang, Nurul I. Rasli, and Hatijah Basri

Abstract This paper presents PSF-ZnO-TiO₂ NPs as SA for passively Q-switched EDFL generation at 1568.6 nm with 3-dB bandwidth of 1 nm. The SA was prepared by cutting the PSF-ZnO-TiO₂ NPs into small pieces and inserting it between two fiber ferrules connectors. The modulation depth of PSF-ZnO-TiO₂ NPs was 45.12% with saturation intensity of 0.0067 MW/cm². The Q-switched pulses generate repetition rate of 19.86 kHz to 50.90 kHz, reducing pulse width from 8.73 to 3.82 μ s from pump power of 194.5 mW to 423.0 mW, respectively. The highest repetition rate was 50.90 kHz, which indicates a SNR value of 34.90 dB. This result shows that PSF-ZnO-TiO₂ NPs has achieved good performance of generating passively Q-switched EDFL, since it is the first demonstration of the use of membrane materials.

1.1 Introduction

Q-switched fiber laser has received many researchers' attention in recent years due to the various potential uses such as in clinical application, remote sensing, optical communication and material processing [1–4]. Compared to the continuous wave laser, the Q-switched fiber laser can generate microseconds to nanoseconds pulse range with repetition rate in the kilohertz regime. Active and passive are techniques to achieve Q-switched pulsed laser. Passive technique is more favorable than active technique due to simple in fabrication, cost-effectiveness and flexibility for industrial applications [5]. This technique can be accomplished via employing a SA into the

N. N. H. E. N. Mahmud · N. A. Awang (✉) · N. I. Rasli · H. Basri
Optical Fiber Laser Technology (OpFLAT) Focus Group, Department of Physics and Chemistry,
Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, 84600
Pagoh, Johor, Malaysia
e-mail: norazura@uthm.edu.my

N. I. Rasli
e-mail: hw170040@siswa.uthm.edu.my

H. Basri
e-mail: hatijah@uthm.edu.my

laser cavity. The first the Q-switched laser is conquered by semiconductor saturable absorber mirrors (SESAMs). Although SESAMs have good stability, but it also suffer from narrow operating wavelengths, cost-ineffective and complex fabrication processes [6, 7].

Other's material is developed and incorporated as SA including graphene [8, 9], carbon nanotubes (CNTs) [10], transition metal dichalcogenides (TMDs) [11–13], black phosphorus [14] and Topological insulator (TI) [15]. Graphene and CNTs have garnered great interest as SAs in Q-switched fiber laser, due to its exceptional broadband optical absorption and fast recovery time [16]. However, graphene has the disadvantages of low modulation depth and weak absorption. At the same time CNTs SA operation wavelength depends on their diameter and can only operate at moderate pumping level as it is relatively low saturation and low damage threshold [17–19].

Over the last few years, two dimensional (2D) materials attracted much attention due to the photonic properties. The 2D material may share similar photonic properties such as various linear optical absorption, causing saturated absorption due to Pauli blocking including fast relaxation time and high optical nonlinearity. Besides graphene, there are three types of 2D materials that are actively investigated which are transition metal dichalcogenides (TMDs), topological insulator (TI) and black phosphorus (BP). The generation of pulsed fiber laser based on topological insulators such as Bi_2Te_3 has been achieved due to its characteristic which is easily saturated beyond a certain threshold. Not only that, the band structure of TIs is similar to graphene with Dirac cones characteristics and offers broader light absorption than graphene [15, 20].

Other than TIs, transition metal oxides (TMOs) are successfully exploited as SA. The transition metal of oxides such as nickel (Ni), titanium (Ti), zinc (Zn), vanadium (V), copper (Cu), aluminum (Al) and Iron (Fe) [21–28]. In 1960, ZnO has gained attention for its various applications such as sensors, transducers and optoelectronic devices due to its 3.37 eV bandgap and high energy binding 60 meV [29] ZnO also has short recovery time of 1–5 ps when observed through second-harmonic generation [30]. Recently, TiO_2 materials are reported to be suitable for pulsed laser operation due to its properties of ~3.2 eV indirect bandgap [31]. Zhang and Wu reported that the spectral absorption of TiO_2 SA material can be extended until the near-infrared (NIR) region and it also can operate in several wavelengths, particularly 1, 1.5 and near 2 μm [22, 32–35]. Moreover, TiO_2 offers fast recovery time of 1.5 ps, cost-effectiveness, easy availability as well as high chemical and optical stability, however it suffers from various instability problems at room temperature operation [36]. Thus, this quality makes it an ideal material for a variety of applications such as water treatment, agriculture and laser application [37–39].

BP also has similar characteristics to graphene, where BP lies between gapless graphene and large bandgap of TMDs. Additionally, this material provides interesting characteristics in terms of direct band gap, planar characteristic and wide-band absorption [40, 41]. Nevertheless, it also encounters some problems when exposed to a humid environment, BP hydrophilic surfaces degrade [42]. According to Song et. al, few solutions to enhance the antioxidant properties onsequently this method makes

the whole process more complicated [43]. Regarding this issue, researchers started to discover material-polymer composite film as SA to improve material stability. This is associated with the hydrophobic properties of polymers that provide encapsulation to the material and ensure oxygen free condition. Guo et al. reported the ultrafast pulse generation using Te/polyvinylpyrrolidone (PVP) composite SA where this material can survive after 6 h stability test in different gain medium. Thus, showing that the composite film method can prolong Te material [44].

In recent years, Polysulfone (PSF) has been the most favorable membrane material due to its desirable thermal and mechanical properties as well as physical and chemical stability [45]. In this work, we demonstrate the fabrication of PSF-ZnO-TiO₂ as SA. Incorporating TiO₂ and ZnO nanoparticles (NPs) in PSF membrane to improve the necessary membrane properties. Among the nanoparticles (NPs) study, TiO₂ NPs has been widely used in preparing nanocomposite membranes due to their good properties such as low-cost, strong photocatalytic, chemical and photochemical stability. However, large band gaps of TiO₂, high electron hole recombination and the aggregation of nanoparticles in membrane matrix and degradation of polymer membranes under UV light. Thus, the combination of TiO₂ with ZnO can lower the rate of electron hole pair recombination, indirectly improving photocatalytic [46, 47]. To our knowledge, this is the first demonstration using membrane material (PSF-ZnO-TiO₂ NPs) as SA to produce a Q-switching pulse. The SA is constructed by sandwiching a PSF-ZnO-TiO₂ NPs between two fiber ferrules. Our finding suggests that PSF-ZnO-TiO₂ NPs could find potential applications in photonic applications.

1.2 Preparation and Characterization of PSF-ZnO-TiO₂ NPs Membrane as Saturable Absorber

1.2.1 Morphology and EDX, Analysis of PSF-ZnO-TiO₂ NPs Membrane

The PSF-ZnO-TiO₂ NPs were synthesized via the phase inversion method. The material preparation consists of 11% of the polymer, 88% of N-methyl-pyrrolidinone (NMP) and 1% of biosynthesized ZnO-TiO₂ NPs. The dope solution was prepared by stirring the polymer, solvent and NPs at 650 rpm and 250 °C overnight. Compared to conventional chemical synthesis methods, the green synthesis route is safer for preparation of biosynthesized NPs. The novel biosynthesized ZnO-TiO₂ NPs. NPs was organized using 2 g of aloe vera extract reacted with 0.3 M of zinc nitrate and 0.3 M of titanium butoxide precursor. The process took 48 h of reaction time at room temperature. The yield of this biosynthesis process was 12.068 g of ZnO-TiO₂ NPs.

The morphology surface of the PSF-ZnO-TiO₂ NPs membrane was characterized by using FESEM (FEI 3D DUAL BEAM) as shown in Fig. 1.1. The images of FESEM were obtained when the electron accelerating voltage was set at 20_{kV} with 50 k magnification, showing a nonuniform shape such as hexagonal, oval and spherical

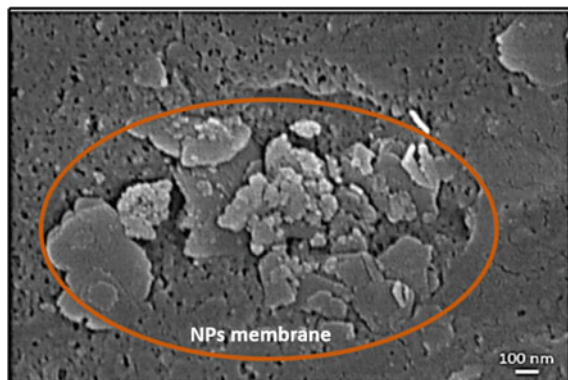


Fig. 1.1 The top surface of PSF-ZnO-TiO₂ NPs at 50kX magnification

due to the mixture of ZnO-TiO₂ NPs [48–51]. The nonuniform shape (in red circle line) measured the NPs size range from 27 to 690 nm at 50kX magnification. The particles also scattered and agglomerated on the top surface of the membrane. The surface morphology for ZnO-TiO₂ NPs shows different structure or shape due to the biosynthesis process. Inspired by the previous research, the NPs are synthesized by using plant extracts which is aloe vera extract [52]. Apart from that, NPs structure can be increased by increasing the concentration of the aloe vera extract.

Figure 1.2 shows the EDX analysis which was operated to determine the elemental of PSF-ZnO-TiO₂ NPs. The EDX analysis examines that carbon (C) has a very strong peak intensity compared to other elements due to the presence of elements shown in the chemical structure of PSF membrane in Fig. 1.3. The existence of other elements such as Titanium (Ti) and Zinc (Zn) also can be seen because of the embedded biosynthesis ZnO-TiO₂ NPs. The respective atomic percentages of each traced element were demonstrated in Table 1.1. The atomic % of the elements shows

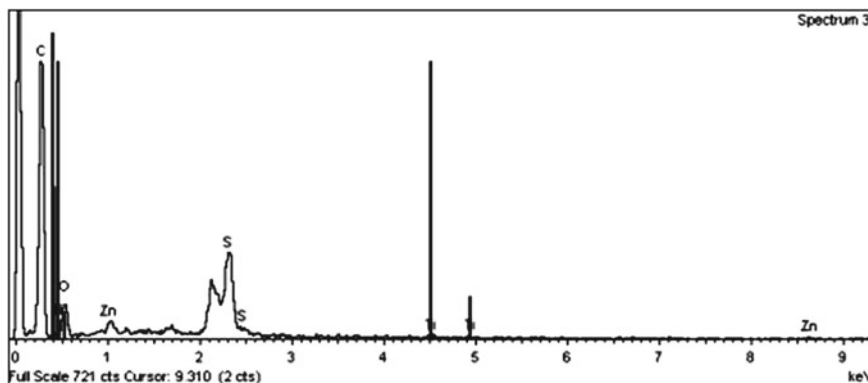


Fig. 1.2 EDX spectrum of PSF-ZnO-TiO₂ NPs

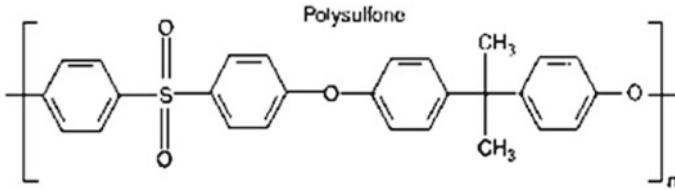


Fig. 1.3 EDX spectrum of PSF-ZnO-TiO₂ NPs

Table 1.1 EDX analysis of elements and atomic %

Elements	Atomic %
C	75.87
O	14.49
S	8.27
Ti	0.14
Zn	1.23

that the number of atoms for C, O, S, Ti and Zn were 75.87%, 14.49%, 8.27%, 0.14% and 1.23% respectively.

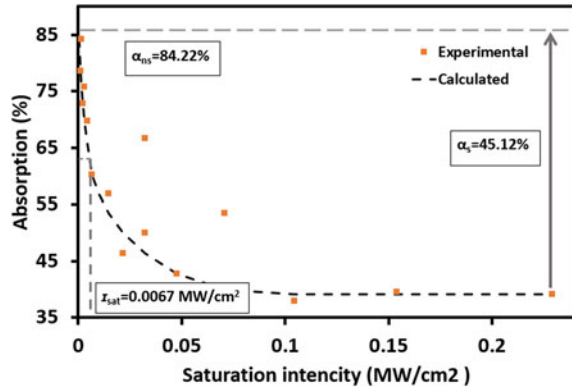
1.2.2 Nonlinear Absorption Properties of PSF-ZnO-TiO₂ NPs SA

The nonlinear optics absorption of PSF-ZnO-TiO₂ NPs SA was studied by the using of the twin detector method. The passive mode-locked EDFL serves as the probe light source with a frequency of 12.64 MHz and pulse width 716.5 fs. Figure 1.3 shows the nonlinear absorption curve. The orange points represent the experimental data, exhibiting a typical saturable absorption performance. The absorption intensity is represented by calculated curve using the formula given by

$$\alpha(I) = \frac{\alpha_s}{1 + \frac{I}{I_s}} + \alpha_{ns} \quad (1.1)$$

Based on the equation above, the absorption coefficient is denoted as $\alpha(I)$. Figure 1.4 displays a saturable absorption (α_s) known as modulation depth, and saturation intensity (I_s) of PSF-ZnO-TiO₂ NPs saturates from absorption data depending on the respective power of 45.12% and 0.0067 MW/cm². Nonlinear absorption can be seen from 39.10 to 84.22% in saturation. This indicates that optical absorption increased by 45.12% when PSF-ZnO-TiO₂ NPs were saturated at an intensity

Fig. 1.4 The nonlinear absorption PSF-ZnO-TiO₂ NPs SA



of 0.0067 MW/cm², giving a modulation depth of 45.12%. Meanwhile, the non-saturable absorption is 84.22% which is stated as a percentage of linear absorption due to normalization [53].

1.3 Experimental Setup of Passively Q-Switched Based on PSF-ZnO-TiO₂ NPs SA

1.3.1 Single Ring Cavity

The prepared PSF-ZnO-TiO₂ NPs SA is incorporated into the laser cavity as presented in Fig. 1.5. From Fig. 1.6, the laser diode 980 nm was pumped through a WDM of 980/1550. The 3 m of EDF with 18.93 dB/m absorption coefficient was employed inside the laser cavity as an amplifying signal. The optical isolator was added right after the EDF to prevent any unidirectional light propagation. Then, PSF-ZnO-TiO₂ NPs SA was cut into a smaller size (less than 1 mm²) and was used into the EDFL cavity by inserting it between two fibers. Once the SA was successfully

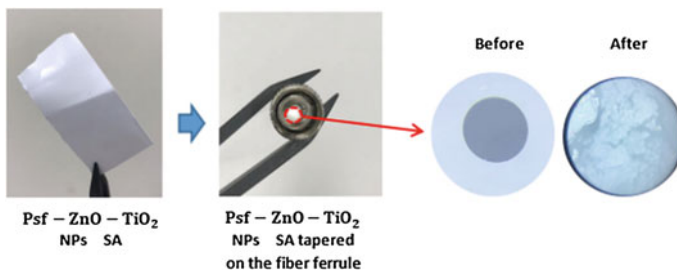


Fig. 1.5 Preparation of PSF-ZnO-TiO₂ NPs SA

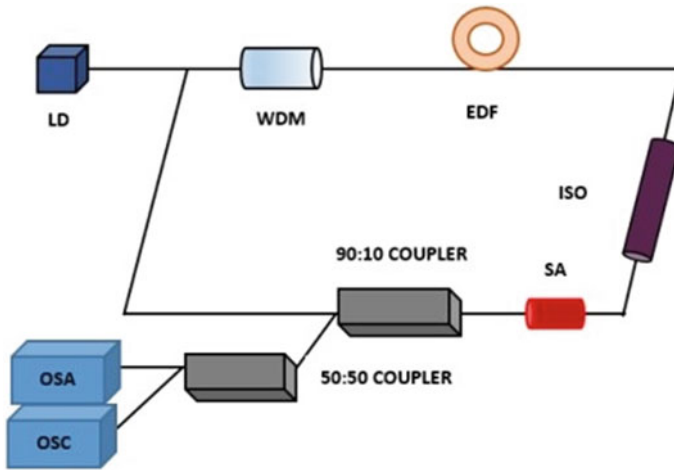


Fig. 1.6 The single ring cavity of passively Q-switched EDFL on PSF-ZnO-TiO₂ NPs SA

employed, the insertion loss of fiber SA was calculated to be less than 1 dB. The 90:10 of output coupler was also added in the laser cavity, where 90% of the light oscillating back inside the cavity. Meanwhile, 10% of light was connected to the input of a 50:50 optical coupler. Signals from the 50:50 coupler were recorded using an Tetronik oscilloscope (OSC) and Anritsu optical spectrum analyzer (OSA). After the single ring cavity was done splicing, the length of the cavity was measured to be 7.44 m, whereas all fiber uses are based on singlemode operation.

1.4 Result and Discussions

Initially, the PSF-ZnO-TiO₂ NPs SA was integrated into the EDFL configuration by sandwiching it between two fiber ferrules. The light signal from the laser diode was pumped at 55.1 mW of pump power to generate a stable CW. The laser cavity reached the threshold pump power at 194.5 mW and the Q-switched pulses started to appear. The pump power kept increasing from 194.1 until the pump power exceeded the maximum pump power of 423.0 mW. Thus, the Q-switched pulses disappeared and laser operation of CW was observed again due to the excessive saturation of SA at high pump power intensity. It proves that the value of repetition rate depends on the pump power applied.

From Fig. 1.7, the optical spectrum of Q-switched EDFL generates the output power of -3.36 dBm. It also shows the central wavelength of the Q-switched pulse in the 1568.6 nm region. Moreover, the laser spectrum operates at 1568.0 and 1569.0 nm with a 3 dB bandwidth of 1 nm which was considered as broad laser bandwidth in

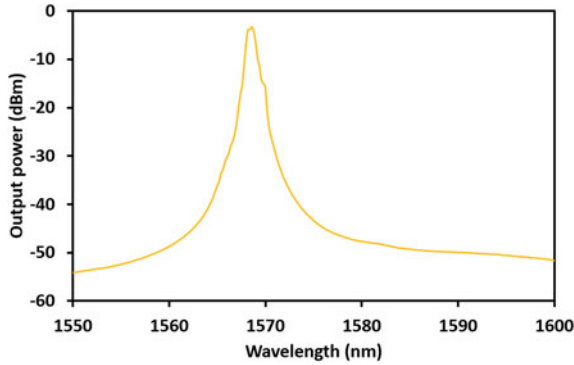


Fig. 1.7 Optical spectrum of passively Q-switched EDFL based on PSF-ZnO-TiO₂ NPs SA at pump power of 423.0 mW

Q-switched spectrum. This broadening effect of the passively Q-switched EDFL occurs because of the self-phase modulation (SPM) phenomena [48].

Figure 1.8a shows the pulse train of the passively Q-switched EDFL. Passively Q-switched EDFL pulse repetition rate was measured at 19.86 kHz at 194.5 mW pump power. The pulse duration measured directly from the oscilloscope was approximately 19.65 μ s. This is due to the self-mode-locking that does not happen during the Q-switched process [49]. Next, the single pulse shape encircled at a minimum pump power of 194.5 mW as shown in Fig. 1.8b. The graph shows the pulse width of the Q-switched EDFL is 8.73 μ s, which corresponds to an output voltage of 0.0016 V. The peak intensity was almost constant at -0.005 and 0.016 V at minimum pump power. This proves that as the pump power increases, the peak intensity remains uniform.

Figure 1.9a shows the pulse train output at the maximum pump power of 423.0 mW. The pulse duration generates 19.65 μ s corresponding to a repetition rate of

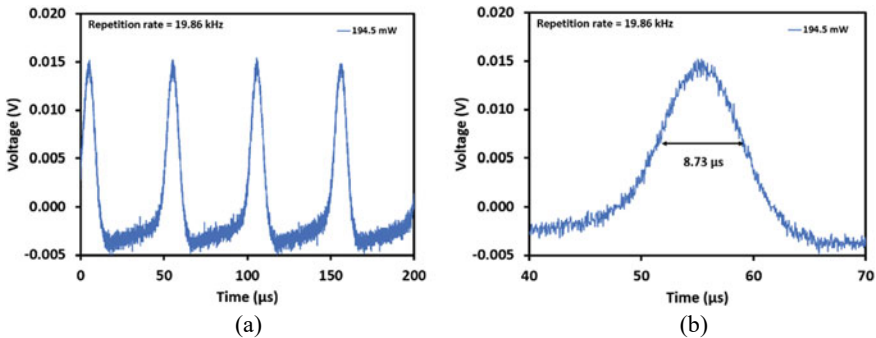


Fig. 1.8 Passively Q-switched characteristic at pump of 194.5 mW **a** pulse train, **b** single pulse profile

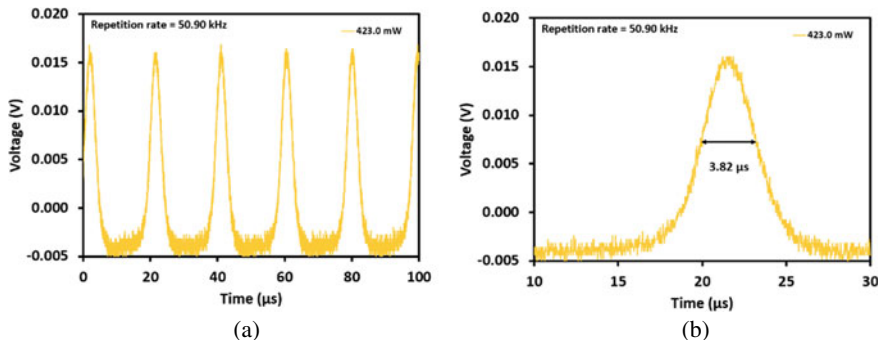


Fig. 1.9 Passively Q-switched characteristic at pump of 423.0 mW **a** pulse train, **b** single pulse profile

50.90 kHz. It shows that the pulse duration depends on the pump power. Apart from that, more gains from the pump power was required to saturate the SA, indicating a narrower pulse width. Figure 1.9b illustrates a single pulse profile which generates a pulse width of 3.82 μ s. The higher pump power applied, the repetition rate increases due to the population inversion, thus giving a narrower pulse width.

To prove the stability of passively Q-switched EDFL performance, the RF spectrum was measured using RF spectrum analyzer with a 5 GHz photodetector. Figure 1.10a, b shows the measured spectrum at minimum and maximum pump power, which indicated the fundamental frequency at 19.86 kHz and 50.90 kHz, respectively. The fundamental frequency of the Q-switched EDFL generates SNR value of 34.30 dB at a minimum pump power of 194.5 mW. Nevertheless, the maximum pump power of 423.0 mW produces SNR value of 34.90 dB which is higher than minimum pump power. Thus, the SNR value increases with pump power and passively Q-switched EDFL shows more stability when more pump power is applied.

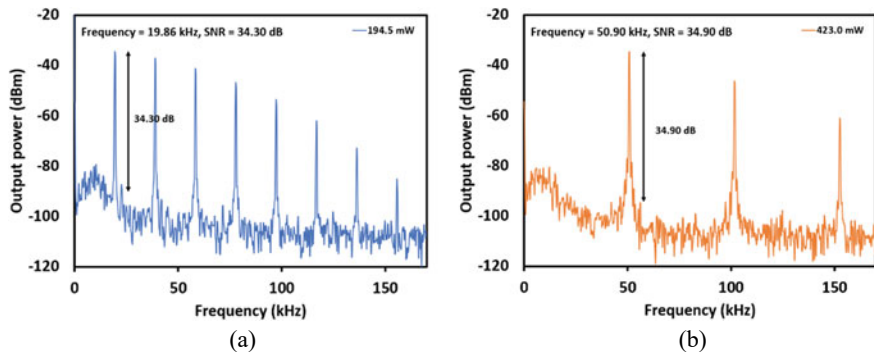
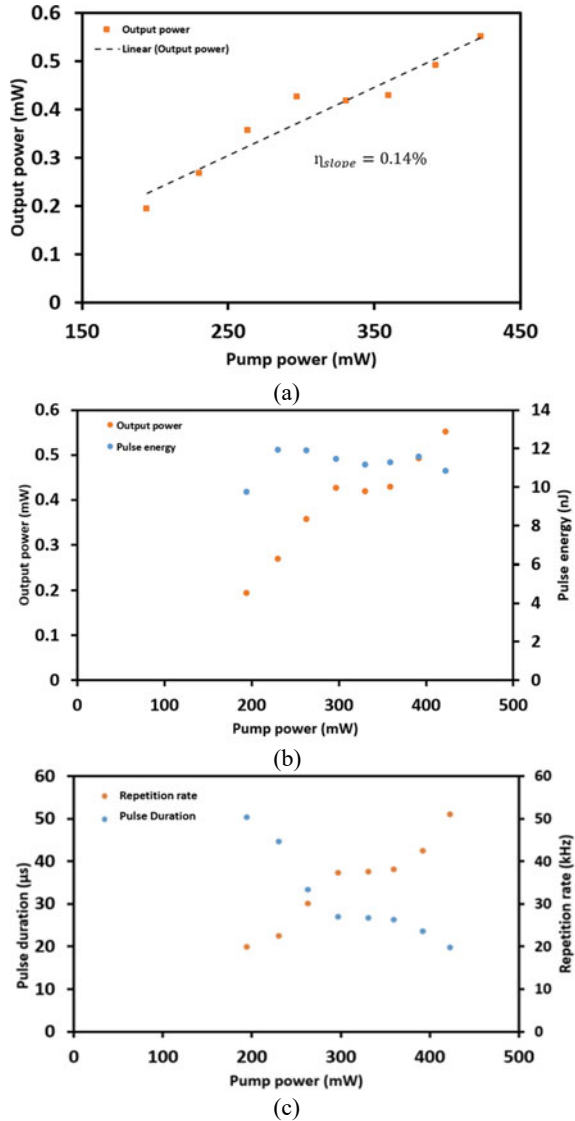


Fig. 1.10 The RF spectrum **a** 194.5 mW, **b** 423.0 mW

Figure 1.11a shows the graph of pump power and output power. The graph shows a linear increment in the output power. From the graph, the slope efficiency was calculated to be 14%. The relationship between pulse energy and output power against pump power shown in Fig. 1.11b. The orange dot indicates that the pulse energy slightly increased and decreased when the pump power reached at maximum. This is due to the unstable Q-switched average output power which can affect the calculation

Fig. 1.11 (a) The output power against pump power (b) the pulse energy and output against pump power (c)



of the energy pulse as the average output power decreases. Therefore, the pulse energy produced at pump power of 230.6 mW was 11.56 nJ which is higher than the others.

Figure 1.11c shows the relationship between pulse duration and repetition rate were measured as a persistence of the pump power. As the pump power increased from 194.5 mW to 423.0 mW, the repetition rate increased linearly from 19.86 to 50.90 kHz, which corresponds to the Q-switched pulse characteristic. As expected, the decrease of pulse duration from 50.35 to 19.65 μ s. Thus, the shortest pulse duration of 19.65 μ s was attained at the pump power of 423.0 mW. The Q-switched pulses become unstable and disappear when the pump power is raised beyond the maximum pump power. This is due to the thermal effect of SA, as the material already saturated.

1.5 Conclusion

The employment of PSF-ZnO-TiO₂ NPs SA for generating passively Q-switched EDFL was successfully demonstrated. PSF-ZnO-TiO₂ NPs SA was fabricated and prepared to generate passive Q-switched, operating at 1568.6 nm with threshold pump power of 194.5 mW. The increase of the pump power from 194.5 mW to 423.0 mW, resulted in the reduction of the pulse duration from 50.35 μ s to 19.65 μ s, which corresponded to the pulse repetitions of 19.86 kHz and 50.90 kHz, respectively. The RF spectrum showed SNR value of about 34.30 dB and 34.90 dB at pump power of 194.5 mW and 423.0 mW. This indicates that the passively Q-switched EDFL operates with good stability. In contrast, the pulse width was reduced from 8.73 μ s to 3.82 μ s when the pump power increased. The proposed Q-switched EDFL incorporating PSF-ZnO-TiO₂ NPs SA showed good performance from the Q-switched pulse train in the 1.5 μ m wavelength region, although this was the first time a PSF membrane embedded with ZnO-TiO₂ NPs biosynthesis. Future work will be focused on the D-shaped fiber incorporating with PSF-ZnO-TiO₂ NPs membrane for pulsed fiber generation.

Acknowledgements The authors would like to thanks the Ministry of Higher Education (MOHE) and for the Fundamental Research Grant Scheme (FRGS) (FRGS/1/2019/WAB05/UTHM/02/4) and Universiti Tun Hussein Onn Malaysia (UTHM) through FRGS (vot K170) for supporting this research through FRGS.

References

1. S.G. Ho, C.L. Goh, Laser tattoo removal: a clinical update. *J. Cutan. Aesthet. Surg.* **8**(1), 9 (2015). <https://doi.org/10.4103/0974-2077.155066>
2. N.P. Barnes, B.M. Walsh, D.J. Reichle, R.J. DeYoung, Tm: fiber lasers for remote sensing. *Opt. Mater.* **31**(7), 1061–1064 (2009)

3. M.J. O'Mahony, D. Simeonidou, D.K. Hunter, A. Tzanakaki, The application of optical packet switching in future communication networks. *IEEE Commun. Mag.* **39**(3), 128–135 (2001)
4. T. Westphäling, Pulsed Fiber Lasers from ns to ms range and their applications. *Phys. Procedia* **5**, 125–136 (2010)
5. K. Wu, X. Zhang, J. Wang, X. Li, J. Chen, WS 2 as a saturable absorber for ultrafast photonic applications of mode-locked and Q-switched lasers. *Opt. Express* **23**(9), 11453–11461 (2015)
6. S.Y. Set, H. Yaguchi, Y. Tanaka, M. Jablonski, Ultrafast fiber pulsed lasers incorporating carbon nanotubes. *IEEE J. Sel. Top. Quantum Electron.* **10**(1), 137–146 (2004)
7. O. Okhotnikov, A. Grudinin, M. Pessa, Ultra-fast fibre laser systems based on SESAM technology: new horizons and applications. *New J. Phys.* **6**(1), 177 (2004)
8. Y.K. Yap, R.M.D.L. Rue, C.H. Pua, S.W. Harun, H. Ahmad, Graphene-based Qswitched pulsed fiber laser in a linear configuration. *Chin. Opt. Lett.* **10**, 041405–041408 (2012). <https://doi.org/10.3788/col201210.041405>
9. D.P. Zhou, L. Wei, W.K. Liu, Tunable graphene Q-switched erbium-doped fiber laser with suppressed self-mode locking effect. *Appl. Opt.* **51**(14), 2554–2558 (2012)
10. B. Dong, J. Hao, J. Hu, C.Y. Liaw, Wide pulse-repetition-rate range tunable nanotubes Q-switched low threshold erbium-doped fiber laser. *IEEE Photonics Technol. Lett.* **22**(24), 1853–1855 (2010)
11. H. Ahmad, M.Z. Samion, A. Muhamad, A.S. Sharbirin, R.A. Shaharuddin, K. Thambiratnam et al., Tunable 2.0 μm Q-switched fiber laser using a silver nanoparticle based saturable absorber. *Laser Phys.* **27**(6), 065110 (2017)
12. J. Sotor, G. Sobon, W. Macherzynski, P. Paletko, K. Grodecki, K.M. Abramski, Mode-locking in Er-doped fiber laser based on mechanically exfoliated Sb 2 Te 3 saturable absorber. *Opt. Mater. Express* **4**(1), 1–6 (2014) (D.P. Zhou, L. Wei, W.K. Liu, Tunable graphene Q-switched erbium-doped fiber laser with suppressed self-mode locking effect. *Appl. Opt.* **51**(14), 2554–2558 (2012))
13. H. Hou, T. You, Q. Zhou, M. Liu, Y. Ouyang, X. Liu, W. Liu, Q-switched all-fiber laser based on titanium trisulfide, *Optik* **205**, 164234 (2020)
14. J. Liu, Y. Chen, Y. Li, H. Zhang, S. Zheng, S. Xu, Switchable dual-wavelength Q-switched fiber laser using multilayer black phosphorus as a saturable absorber. *Photonics Res.* **6**(3), 198–203 (2018)
15. C. Zhao, H. Zhang, X. Qi, Y. Chen, Z. Wang, S. Wen, D. Tang, Ultra-short pulse generation by a topological insulator based saturable absorber. *Appl. Phys. Lett.* **101**(21), 211106 (2012)
16. A. Martinez, Z. Sun, Nanotube and graphene saturable absorbers for fibre lasers. *Nat. Photonics* **7**(11), 842–845 (2013)
17. S. Yamashita, Y. Inoue, S. Maruyama, Y. Murakami, H. Yaguchi, M. Jablonski, S.Y. Set, Saturable absorbers incorporating carbon nanotubes directly synthesized onto substrates and fibers and their application to mode-locked fiber lasers. *Opt. Lett.* **29**(14), 1581–1583 (2004)
18. A. Schmidt, S. Rivier, G. Steinmeyer, J.H. Yim, W.B. Cho, S. Lee et al., Passive mode locking of Yb: KLuW using a single-walled carbon nanotube saturable absorber. *Opt. Lett.* **33**(7), 729–731 (2008)
19. J. Sotor, G. Sobon, K. Krzempek, K.M. Abramski, Fundamental and harmonic mode-locking in erbium-doped fiber laser based on graphene saturable absorber. *Opt. Commun.* **285**(13–14), 3174–3178 (2012)
20. H. Haris, H. Arof, A.R. Muhammad, C.L. Anyi, S.J. Tan, N. Kasim, S.W. Harun, Passively Q-switched and mode-locked Erbium-doped fiber laser with topological insulator Bismuth Selenide (Bi₂Se₃) as saturable absorber at C-band region. *Opt. Fiber Technol.* **48**, 117–122 (2019)
21. I.A.M. Alani, B.A. Ahmad, M.H.M. Ahmed, A.A. Latiff, A.H.H. Al-Masoodi, M.Q. Lokman, S.W. Harun, Nanosecond mode-locked erbium doped fiber laser based on zinc oxide thin film saturable absorber. *Indian J. Phys.* **93**(1), 93–99 (2019)
22. N. Hisamuddin, Z. Jusoh, U.N. Zakaria, M.Z. Zulkifli, A.A. Latiff, M. Yasin et al., Q-switched Raman fiber laser with titanium dioxide based saturable absorber. *Optoelectron. Adv. Mater. Rapid Commun.* **11**(March-April 2017), 127–130 (2017)

23. A. Nady, M.H.M. Ahmed, A.A. Latiff, A. Numan, C.R. Ooi, S.W. Harun, Nickel oxide nanoparticles as a saturable absorber for an all-fiber passively Q-switched erbium-doped fiber laser. *Laser Phys.* **27**(6), 065105 (2017)
24. A. Nady, M.F. Baharom, A.A. Latiff, S.W. Harun, Mode-locked erbium-doped fiber laser using vanadium oxide as saturable absorber. *Chin. Phys. Lett.* **35**(4), 044204 (2018)
25. S.A. Sadeq, S.W. Harun, A.H. Al-Janabi, Ultrashort pulse generation with an erbium-doped fiber laser ring cavity based on a copper oxide saturable absorber. *Appl. Opt.* **57**(18), 5180–5185 (2018)
26. S.K.M. Al-Hayali, D.Z. Mohammed, W.A. Khaleel, A.H. Al-Janabi, Aluminum oxide nanoparticles as saturable absorber for C-band passively Q-switched fiber laser. *Appl. Opt.* **56**(16), 4720–4726 (2017)
27. J. Koo, J. Lee, J. Kim, J.H. Lee, A Q-switched, 1.89 μm fiber laser using an Fe_3O_4 -based saturable absorber. *J. Lumin.* **195**, 181–186 (2018)
28. D. Mao, X. Cui, W. Zhang, M. Li, T. Feng, B. Du et al., Q-switched fiber laser based on saturable absorption of ferroferric-oxide nanoparticles. *Photonics Res.* **5**(1), 52–56 (2017)
29. H. Ahmad, C.S.J. Lee, M.A. Ismail, Z.A. Ali, S.A. Reduan, N.E. Ruslan et al., Zinc oxide (ZnO) nanoparticles as saturable absorber in passively Q-switched fiber laser. *Opt. Commun.* **381**, 72–76 (2016)
30. J.C. Johnson, K.P. Knutsen, H. Yan, M. Law, Y. Zhang, P. Yang, R.J. Saykally, Ultrafast carrier dynamics in single ZnO nanowire and nanoribbon lasers. *Nano Lett.* **4**(2), 197–204 (2004)
31. K.M. Reddy, S.V. Manorama, A.R. Reddy, Bandgap studies on anatase titanium dioxide nanoparticles. *Mater. Chem. Phys.* **78**(1), 239–245 (2003)
32. Z. Zhang, H. Wu, Multiple band light trapping in ultraviolet, visible and near infrared regions with TiO_2 based photonic materials. *Chem. Commun.* **50**(91), 14179–14182 (2014)
33. M.F.A. Rahman, P.H. Reddy, M.C. Paul, S. Das, A. Dhar, M.F. Baharom et al., Titanium dioxide fiber saturable absorber for Q-switched fiber laser generation in the 1-micrometer region. *Appl. Opt.* **58**(13), 3495–3500 (2019)
34. M.F.M. Rusdi, A.A. Latiff, M.C. Paul, S. Das, A. Dhar, H. Ahmad, S.W. Harun, Titanium dioxide (TiO_2) film as a new saturable absorber for generating modelocked Thulium-Holmium doped all-fiber laser. *Opt. Laser Technol.* **89**, 16–20 (2017)
35. B. Karunakaran, K. Kim, D. Mangalaraj, J. Yi, S. Velumani, Structural, optical and Raman scattering studies on DC magnetron sputtered titanium dioxide thin films. *Sol. Energy Mater. Sol. Cells* **88**(2), 199–208 (2005)
36. C. Luo, X. Ren, Z. Dai, Y. Zhang, X. Qi, C. Pan, Present perspectives of advanced characterization techniques in TiO_2 -based photocatalysts. *ACS Appl. Mater. Interfaces* **9**(28), 23265–23286 (2017)
37. W.M. Mahmoud, T. Rastogi, K. Kümmerer, Application of titanium dioxide nanoparticles as a photocatalyst for the removal of micropollutants such as pharmaceuticals from water. *Curr. Opin. Green Sustain. Chem.* **6**, 1–10 (2017)
38. A.A.H. Abdel Latef, A.K. Srivastava, M.S.A. El-Sadek, M. Kordrostami, L.S.P. Tran, Titanium dioxide nanoparticles improve growth and enhance tolerance of broad bean plants under saline soil conditions. *Land Degrad. Dev.* **29**, 1065–1073 (2018)
39. E.K. Ng, K.Y. Lau, H.K. Lee, M.H.A. Bakar, Y.M. Kamil, M.F. Omar, M.A. Mahdi, Saturable absorber incorporating graphene oxide polymer composite through dip coating for mode-locked fiber laser. *Opt. Mater.* **100** (2020) 109619
40. A.H.H. Al-Masoodi, M.H.M. Ahmed, A.A. Latiff, H. Arof, S.W. Harun, Q-switched ytterbium-doped fiber laser using black phosphorus as saturable absorber. *Chin. Phys. Lett.* **33**(5), 054206 (2016)
41. E.I. Ismail, N.A. Kadir, A.A. Latiff, H. Ahmad, S.W. Harun, Black phosphorus crystal as a saturable absorber for both a Q-switched and mode-locked erbium-doped fiber laser. *RSC Adv.* **6**(76), 72692–72697 (2016)
42. Y. Huang, J. Qiao, K. He, S. Bliznakov, E. Sutter, X. Chen, P. Sutter, Interaction of black phosphorus with oxygen and water. *Chem. Mater.* **28**(22), 8330–8339 (2016)

43. Y. Song, Z. Liang, X. Jiang, Y. Chen, Z. Li, L. Lu, H. Zhang, Few-layer antimonene decorated microfiber: ultra-short pulse generation and all-optical thresholding with enhanced long-term stability. *2D Mater.* **4**(4), 045010 (2017)
44. J. Guo, J. Zhao, D. Huang, Y. Wang, F. Zhang, Y. Ge, H. Zhang, Two-dimensional tellurium-polymer membrane for ultrafast photonics. *Nanoscale* **11**(13), 6235–6242 (2019)
45. B. Zhao, Z. Wang, S. Zhao, J. Wang, S. Wang, Preparation and performance of PSF/PANI-F127 composite ultrafiltration membrane. *CIESC J.* **64**(2), 702–710 (2013)
46. A. Mills, S. Le Hunte, An overview of semiconductor photocatalysis. *J. Photochem. Photobiol. A* **108**(1), 1–35 (1997)
47. H. Zangeneh, A.A. Zinatizadeh, S. Zinadini, M. Feyzi, D.W. Bahnemann, Preparation ultrafine L-Methionine (C, N, S triple doped)-TiO₂-ZnO nanoparticles and their photocatalytic performance for fouling alleviation in PES nanocomposite membrane. *Compos. Part B: Eng.* **176**, 107158 (2019)
48. G. Sangeetha, S. Rajeshwari, R. Venckatesh, Green synthesis of zinc oxide nanoparticles by aloe barbadensis miller leaf extract: structure and optical properties. *Mater. Res. Bull.* **46**(12), 2560–2566 (2011)
49. M.A. Irshad, R. Nawaz, M.Z. ur Rehman, M. Adrees, M. Rizwan, S. Ali, S. Tasleem, Synthesis, characterization and advanced sustainable applications of titanium dioxide nanoparticles: a review. *Ecotoxicol. Environ. Safety* **212**, 111978 (2021)
50. M. Aslam, A.Z. Abdullah, M. Rafatullah, Recent development in the green synthesis of titanium dioxide nanoparticles using plant-based biomolecules for environmental and antimicrobial applications. *J. Ind. Eng. Chem.* (2021)
51. N.I. Rasli, H. Basri, Z. Harun, Zinc oxide from aloe vera extract: two-level factorial screening of biosynthesis parameters. *Heliyon* **6**(1), e03156 (2020)
52. N. I. Rasli, H. Basri, Z. Harun, Biosynthesized nanoparticles from aloe vera: A brief review towards membrane technology. *Malaysian J Fundam and Appl Sci*, **15**(6), 895–902 (2019)
53. Q. Bao, H. Zhang, Z. Ni, Y. Wang, L. Polavarapu, Z. Shen, K.P. Loh, Monolayer graphene as a saturable absorber in a mode-locked laser. *Nano Res.* **4**(3), 297–307 (2011)

Chapter 2

Validation and Thickness Dependency of a Planar Material Return Loss by Numerical Simulation



Wan Azmafaliyana Wan Azmi and Fahmiruddin Esa

Abstract This work aims to validate return loss (RL) calculation based on transmission line (TRL) equation method of graphene oxide zinc ferrite (GO-ZnFe₂O₄) epoxy composite through numerical design simulation using CST-Microwave Studio. Electromagnetic properties of material are important parameter for this work, which have been retrieved earlier based on measured scattering parameter (S-parameter) in previous study. Thus, a two-port rectangular waveguide (rwg) was firstly designed and simulated to validate the S-parameter, reflection coefficient, S_{11} and transmission scattering, S_{21} at X-band frequency and both results are found to be perfectly matched with each other. Second simulation work on planar composite based on one-port termination method which was designed for RL characteristic where two types of configurations were considered; waveguide and free space. The result of minimum RL frequency by free space method is in good agreement with the TRL that has been reported, which is at 10.5 GHz. However, the RL characteristics are distinct between each other due to higher thickness that will cause higher reflection in frequency selective surface (FSS) mode of simulation. The proposed thickness for GO and 3-GO-ZnFe₂O₄ as a planar microwave absorber at X-band frequency range is from 7 to 8 mm due to good impedance matching.

2.1 Introduction

The rapid development of electronic and telecommunications systems has led to a strong interest in the electromagnetic absorber and microwave technology which have been commonly used in various fields, such as stealth technology [1], EM compatibility and shielding, sensors [2], energy harvesting [3], and passive cooling technologies [4].

W. A. W. Azmi · F. Esa (✉)

Department of Physics and Chemistry, Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, 84600 Pagoh, Muar, Johor, Malaysia
e-mail: fahmir@uthm.edu.my

The promising microwave absorbing material can be found in composite materials with dielectric and magnetic fillers due to their excellent dielectric and magnetic characteristics along with a lower density and resistance to chemical substances [5]. A matrix and filler can be found in a microwave absorber material. A polymer is commonly used as a matrix for its low processing temperature and flexibility of design [6]. Commonly, most microwave absorbing materials are made of magnetic loss powders such as magnetic metals [7], ferrites [8] and dielectric loss materials such as carbon black [9], carbon fibers [10], carbon coils [11], carbon nanotubes (CNTs) [12] and silicon carbide fibers [13] and polymers [14]. Ferrites as magnetic fillers are one of the strongest magnetic materials ever found because of their low costs, strong chemical stability, and variability in technological use [15]. The characteristics of graphene-based materials, such as low density, wide surface area and high chemical stability [16] have become a desirable option for microwave absorption materials.

S-parameters are being used to characterize the reaction of any network to the signals existing in its port(s). The S-parameters define the magnitude and phase relation between the incident and the reflected waves on its port [17]. The S-parameter is typically attached to the receiver port and the source port such as S_{ij} where i represent the receiver while j represents the source port.

Meanwhile, RL is a reflecting coefficient denoted by (S_{11}) which represents the inputs and outputs relation among ports and how much power the composite material reflects [18]. When S_{11} becomes 0 dB then it indicates that all power is being reflected from the composite material, nothing is being dissipated. The value of S_{11} should be at least -10 dB to provide effective output and then become a suitable candidate in the application of microwave absorbing material [19]. Equation 2.1 shows the equation of RL characteristics of composite materials by using transmission line theory [20].

$$ReturnLoss(dB) = 20 \log \left| \frac{Z_{in} - 1}{Z_{in} + 1} \right| \quad (2.1)$$

$$Z_{in} = \sqrt{\frac{\mu_r}{\epsilon_r}} \tanh(j \frac{2\pi f d}{c} \sqrt{\mu_r \epsilon_r}) \quad (2.2)$$

where Z_{in} is an input impedance of the absorber (epoxy composite material), ϵ_r is complex relative permittivity, μ_r is the permeability of absorber, f is the frequency of the incident electromagnetic wave, d is sample thickness and c is the velocity of light in the free space.

In this work, by considering the results found from the previous study by [19], the reported data of S-parameter has not been validated via simulation. The electromagnetic properties (permittivity and permeability), were well extracted by Nicolson Ross Weir (NRW) conversion method where the S-parameters were measured using closed rectangular waveguide technique. Thus, it is a good motivation to design a rectangular waveguide and simulate for S-parameters based on the calculated electromagnetic properties. Investigation on RL characteristic by simulation with one termination port also has become an interest in this work where comparison will be

made between free space and waveguide configurations. The finding later probably would give us the result of which configuration fits well with the one that using the TRL equation [19].

2.2 Methodology

CST Microwave Studio software, which is based on finite difference time domain (FDTD) was used to design a two-port rectangular waveguide (rwg) for S-parameter validation. Meanwhile for RL validation basically is created on rectangular waveguide with one port termination and also frequency selective surface (FSS) which are defined as waveguide (WG) and free space (FS) method, respectively in this study.

Figure 2.1a shows the rwg design where the material type defined as perfect electric conductor (PEC). The composite materials of 3-GO-ZnFe₂O₄, 1-GO-ZnFe₂O₄, GO and ZnFe₂O₄ were built inside and at the middle of the rwg length as shown in Fig. 2.1b where the frequency dependent electromagnetic properties were earlier retrieved as reported in [19] with their respective thicknesses. Port 1 and Port 2 were defined at each end of rwg, which represents microwave excitation and reception, respectively.

In this work, dimension of the waveguide was designed based on WR90 size standard where the operating frequency is in X-band range (8.2–12.4 GHz). The most important parameter is the cross-sectional area of the aperture, which is 22.86 mm (w) \times 10.16 mm (h) that would give 6.56 GHz cut-off frequency. The length, l and thickness, t of the rwg were designed 100 mm long and 2 mm thick, respectively.

Two types of RL characteristic simulation were made in order to compare with the one that has been simulated using TRL equation as reported by [19]. Firstly, rwg with

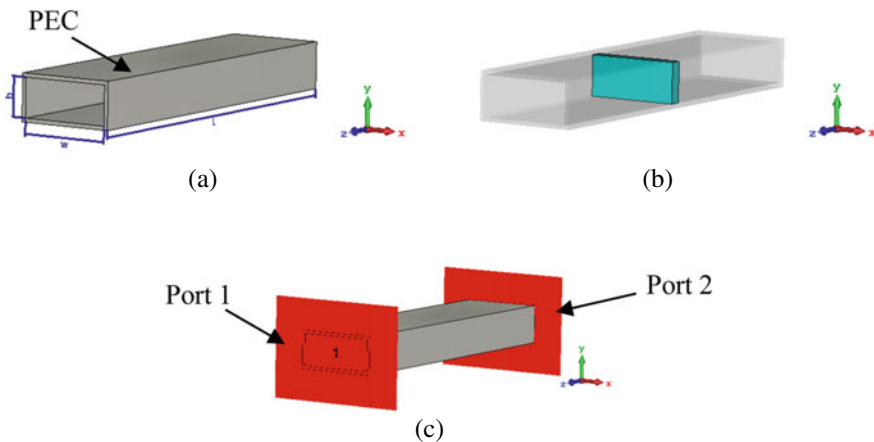


Fig. 2.1 **a** An empty rwg, **b** composite material inside the rwg and **c** full configuration of the two-port rwg

one termination port (WG) at the end of one of the apertures. This design basically the same as rwg as discussed in the previous section in terms of the cross-sectional dimension, however only excitation port was considered and the planar composite materials were placed exactly right after the terminated port as shown in Fig. 2.2. The second design was based on frequency selective surface (FSS) mode of simulation, which is mostly used by periodical structure design of metamaterials. However, only planar design of materials with various thickness will be investigated and because of the measurement and simulation configuration can be done in an open air, thus this technique is assumed to be like free space simulation, (FS). Figure 2.3 shows 8×8 periodical structure of planar composite with 23.7 mm^2 single unit cell.

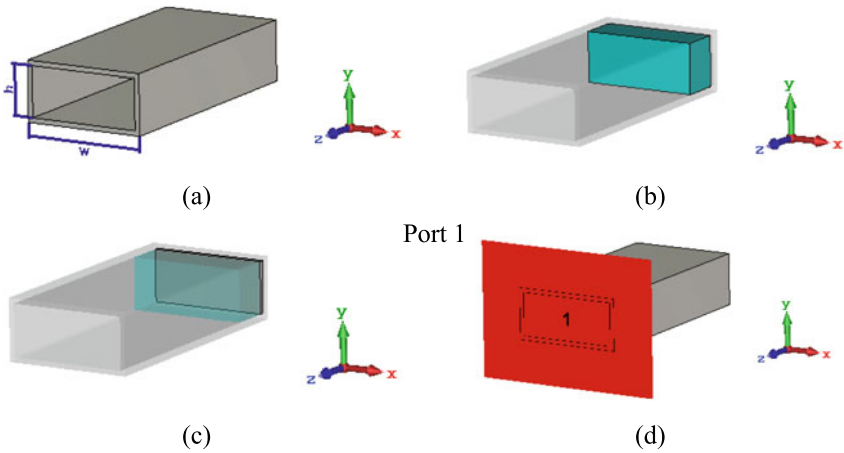


Fig. 2.2 a One-port rwg, b composite material at the terminated port, c termination port and d full configuration of WG

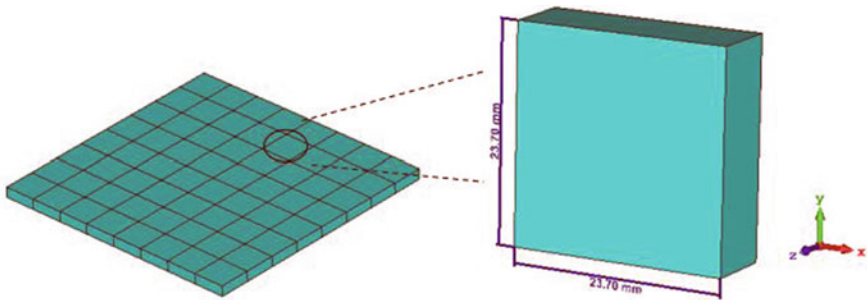


Fig. 2.3 Periodical structure of planar composite. The black circle shaped represents a unit cell with a square dimension of $23.7 \times 23.7 \text{ mm}$ that can be in various thickness

2.3 Result and Discussion

Figure 2.4a, b show the results of S_{11} and S_{21} , respectively at the X-band frequency. It is obviously seen, there is a good agreement between both measured and simulated data. This proves that the retrieved electromagnetic properties of the composite done by [19] were accurate due to the correct measurement of S-parameters using a rectangular waveguide method. The S_{11} is increased as the GO content in the composite increases because GO has a significant reflection as compared to $ZnFe_2O_4$ due to higher conductivity [21]. Whereas S_{21} shows an inverse trending where the transmission is less significant for GO due to higher dielectric loss [19]. Generally, at least two epoxy composites having higher loss mechanisms, which are GO and 3-GO- $ZnFe_2O_4$. These composites are probably the best candidates to be as microwave absorber and further investigation will be discussed in the following subsection.

Figure 2.5 shows the RL characteristic of planar composites, GO and 3-GO- $ZnFe_2O_4$ at low and high thicknesses based on TRL, FS and WG method. The composites with low thicknesses were used for S-parameters measurement and relative permittivity and permeability determination using NRW conversion technique, which is already done in the previous study [19]. In the report also mentioned about increasing the RL characteristic performance by changing composite's thickness while conserve the electromagnetic properties so that optimum matching impedance could be reached via TRL method. Thus, by keeping all those properties and thickness parameters persistent, RL is continuously being explored in this work using FS and WG methods. Apparently, the minimum RL characteristic for both FS and WG method seems to have a large gap between TRL and are distinct from each other especially for the higher thickness of the composite sample. This is because FS method is based on frequency selective surface (FSS) procedure where the suitability for low thickness material such as metamaterial [22]. The effect of higher thickness will cause higher reflection which in return gives a less significant result of minimum RL value. While WG method, the selection of sample thickness must be less than a quarter lambda due to its matching impedance [23]. However, the minimum RL

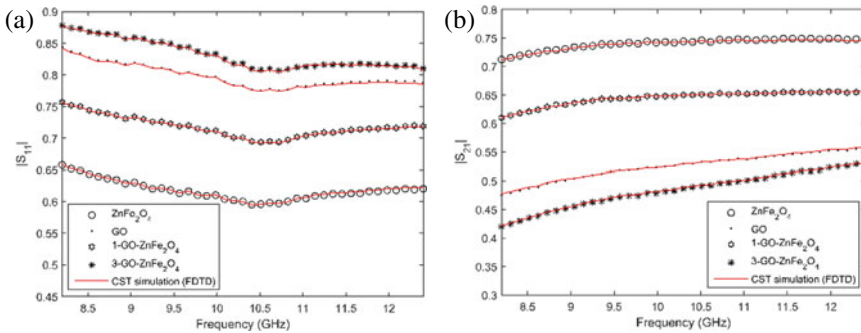


Fig. 2.4 a S_{11} and b S_{21} at X-band frequency

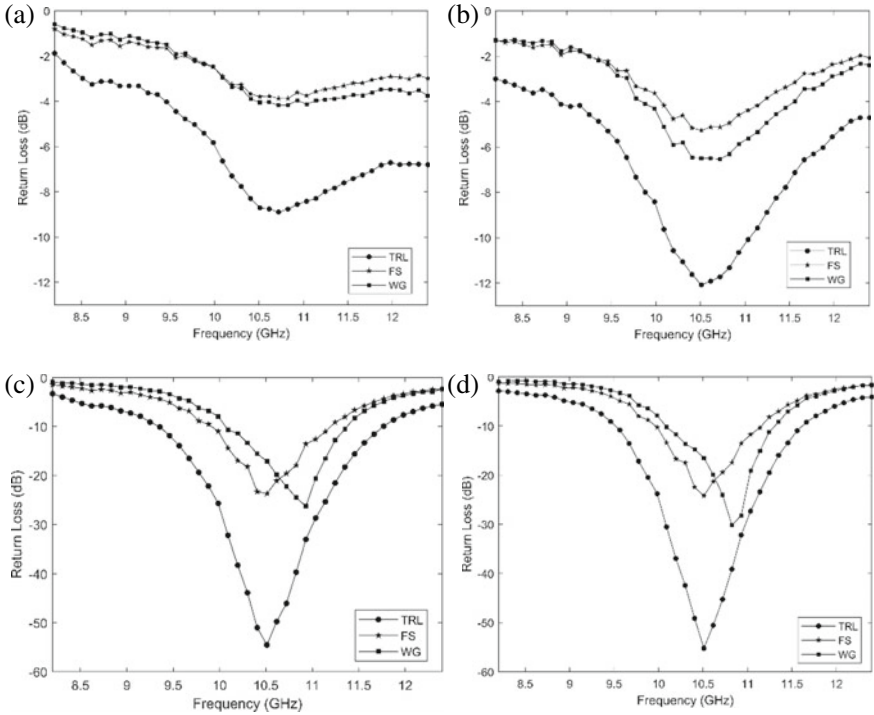


Fig. 2.5 RL characteristic of GO and 3-GO-ZnFe₂O₄ at thickness **a** 2.28 mm, **b** 2.35 mm, **c** 7.8 mm and **d** 7.2 mm, respectively

frequency of FS simulations mostly occurred at a similar frequency of TRL, which is 10.5 GHz that due to the dissipation factor caused by dielectric and magnetic loss. This would suggest that the best simulation result in this study for validating the TRL is by FS method. Table 2.1 shows a summary taken from the simulation results.

Figure 2.6 shows the RL characteristics of GO and 3-GO-ZnFe₂O₄ at various thicknesses from 7 to 8 mm by 0.2 mm increment using FS method. It can be clearly seen that both composites have an ideal matching impedance with thickness of 7.8 mm and 7.2 mm, respectively in the X-band frequency range where the RL values achieved the minimum level at 10.5 GHz. This finding might be supported by the result obtained using the TRL method earlier. Another useful information from this simulation results, the multiple absorption of composite material which is less than -10 dB could occur at different resonant frequency, range and bandwidth depends on application.

Table 2.1 The RL characteristics of planar composite of GO and 3-GO-ZnFe₂O₄

Composite		GO		3-GO-ZnFe ₂ O ₄	
Thickness (mm)		2.28	7.8	2.35	7.20
Method	RL characteristics				
TRL	Resonant frequency (GHz)	10.60	10.50	10.50	10.50
	Min RL (dB)	-8.50	-55.00	-12.50	-55.0
	Range (GHz) at -10 dB	NA	9.40-11.70	10.20-11.00	9.50-11.60
	Bandwidth (GHz)	NA	2.30	0.80	2.10
FS	Resonant frequency (GHz)	10.72	10.51	10.51	10.51
	Min RL (dB)	-3.87	-23.73	-5.27	-24.20
	Range (GHz) at -10 dB	NA	9.99-11.14	NA	9.98-11.14
	Bandwidth (GHz)	NA	1.15	NA	1.16
WG	Resonant frequency (GHz)	10.72	10.93	10.72	10.83
	Min RL (dB)	-4.17	-26.32	-6.54	-30.15
	Range (GHz) at -10 dB	NA	10.09-11.35	NA	10.09-11.25
	Bandwidth (GHz)	NA	1.26	NA	1.16

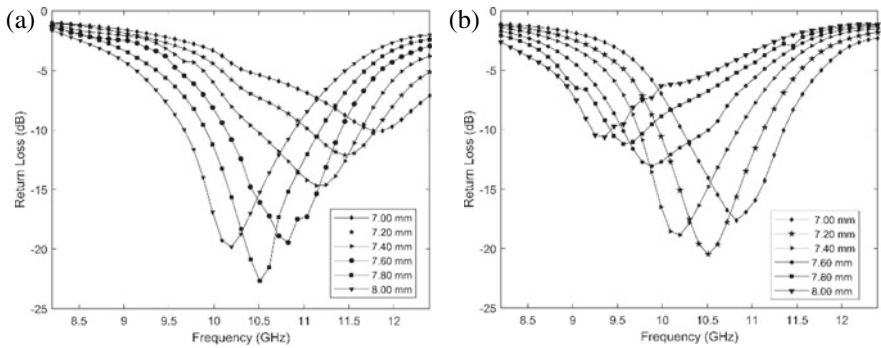


Fig. 2.6 RL characteristic of **a** GO, **b** 3-GO-ZnFe₂O₄ at various thickness

2.4 Conclusion

In conclusion, S-parameters of S_{11} and S_{21} are successfully validated through two-port rwg simulation. FS and WG are proposed in this work to validate RL calculation using TRL method. Resonant frequency of RL simulated using FS method

is very close with TRL, which is 10.5 GHz but the minimum RL, absorption range and bandwidth quite differ due to higher thickness that can cause higher reflection. However, the extended of FS simulation agrees that the best thickness of GO and 3-GO-ZnFe₂O₄ due to good matching impedance is in the range of 7–8 mm.

1. References

1. W. Li, T. Wu, W. Wang, P. Zhai, J. Guan, Broadband patterned magnetic microwave absorber. *J. Appl. Phys.* **116**(4) (2014). <https://doi.org/10.1063/1.4891475>
2. B.X. Wang, X. Zhai, G.Z. Wang, W.Q. Huang, L.L. Wang, A novel dual-band terahertz metamaterial absorber for a sensor application. *J. Appl. Phys.* **117**(1) (2015). <https://doi.org/10.1063/1.4905261>
3. M. Karaaslan, M. Bağmancı, E. Ünal, O. Akgöl, O. Altuntaş, C. Sabah, Broad band metamaterial absorber based on wheel resonators with lumped elements for microwave energy harvesting. *Opt. Quantum Electron.* **50**(5), 1–18 (2018). <https://doi.org/10.1007/s11082-018-1484-2>
4. C. Zou et al., Metal-loaded dielectric resonator metasurfaces for radiative cooling. *Adv. Opt. Mater.* **5**(20), 1–7 (2017). <https://doi.org/10.1002/adom.201700460>
5. H. Zhang, A. Xie, C. Wang, H. Wang, Y. Shen, X. Tian, Novel rGO/ α -Fe₂O₃ composite hydrogel: synthesis, characterization and high performance of electromagnetic wave absorption. *J. Mater. Chem. A* **1**(30), 8547–8552 (2013). <https://doi.org/10.1039/c3ta11278k>
6. H. Gargama, A.K. Thakur, S.K. Chaturvedi, Polyvinylidene fluoride/nanocrystalline iron composite materials for EMI shielding and absorption applications. *J. Alloys Compd.* **654**, 209–215 (2016). <https://doi.org/10.1016/j.jallcom.2015.09.059>
7. Q. Liu et al., Dependency of magnetic microwave absorption on surface architecture of Co₂₀Ni₈₀ hierarchical structures studied by electron holography. *Nanoscale* **7**(5), 1736–1743 (2015). <https://doi.org/10.1039/c4nr05547k>
8. F. Wang, X. Wang, J. Zhu, H. Yang, X. Kong, X. Liu, Lightweight NiFe₂O₄ with controllable 3D network structure and enhanced microwave absorbing properties. *Sci. Rep.* **6**(July), 1–8 (2016). <https://doi.org/10.1038/srep37892>
9. A.A. Al-ghamdi et al., Conductive carbon black/magnetite hybrid fillers in microwave absorbing composites based on natural rubber. *Compos. Part B* **96**, 231–241 (2016). <https://doi.org/10.1016/j.compositesb.2016.04.039>
10. A. Shah et al., Microwave absorption and flexural properties of Fe nanoparticle/carbon fiber/epoxy resin composite plates, vol. 131. Elsevier Ltd (2015)
11. J. Hu, T. Zhao, X. Peng, W. Yang, X. Ji, T. Li, Growth of coiled amorphous carbon nanotube array forest and its electromagnetic wave absorbing properties. *Compos. Part B Eng.* **134**(3), 91–97 (2018). <https://doi.org/10.1016/j.compositesb.2017.09.071>
12. I.W. Nam, H.K. Lee, J.H. Jang, Electromagnetic interference shielding/absorbing characteristics of CNT-embedded epoxy composites. *Compos. Part A Appl. Sci. Manuf.* **42**(9), 1110–1118 (2011). <https://doi.org/10.1016/j.compositesa.2011.04.016>
13. H. Zhang, J. Zhang, H. Zhang, Electromagnetic properties of silicon carbide foams and their composites with silicon dioxide as matrix in X-band. *Compos. Part A Appl. Sci. Manuf.* **38**(2), 602–608 (2007). <https://doi.org/10.1016/j.compositesa.2006.02.014>
14. C. Tian et al., Constructing uniform core-shell PPy@PANI composites with tunable shell thickness toward enhancement in microwave absorption. *ACS Appl. Mater. Interfaces* **7**(36), 20090–20099 (2015). <https://doi.org/10.1021/acsami.5b05259>
15. J. Kruželák, R. Dosoudil, R. Sýkora, I. Hudec, Rubber composites with incorporated magnetic filler. *Polimery/Polymers* **59**(11–12), 819–824 (2014). <https://doi.org/10.14314/polimery.2014.819>
16. X. Zhao et al., Excellent microwave absorption property of Graphene-coated Fe nanocomposites. *Sci. Rep.* **3**, 1–5 (2013). <https://doi.org/10.1038/srep03421>

17. H. Ahmed, J. Hyun, J.R. Lee, Development of scanning single port free space measurement setup for imaging reflection loss of microwave absorbing materials. *Meas. J. Int. Meas. Confed.* **125**, 114–122 (2018) (no. June 2017). <https://doi.org/10.1016/j.measurement.2018.04.065>
18. M. Faisal, A. Gafur, S.Z. Rashid, M.O. Shawon, K.I. Hasan, M.B. Billah, in *Return Loss and Gain Improvement for 5G Wireless Communication Based on Single Band Microstrip Square Patch Antenna. 1st Int. Conf. Adv. Sci. Eng. Robot. Technol. 2019, ICASERT 2019*, vol. 2019 (2019), no. Icasert, pp. 1–5. <https://doi.org/10.1109/ICASERT.2019.8934474>
19. Y.L. Chan et al., Investigation into return loss characteristic of graphene oxide/zinc ferrite/epoxy composite at X-band frequency. **23**(4), 593–602 (2020)
20. T. Maeda, S. Sugimoto, T. Kagotani, N. Tezuka, K. Inomata, Effect of the soft/hard exchange interaction on natural resonance frequency and electromagnetic wave absorption of the rare earth-iron-boron compounds. *J. Magn. Magn. Mater.* **281**(2–3), 195–205 (2004). <https://doi.org/10.1016/j.jmmm.2004.04.105>
21. R. Shu et al., Synthesis and high-performance microwave absorption of reduced graphene oxide/zinc ferrite hybrid nanocomposite. *Mater. Lett.* **215**, 229–232 (2018). <https://doi.org/10.1016/j.matlet.2017.12.108>
22. R. Panwar, S. Puthucheri, D. Singh, V. Agarwala, J.R. Lee, Microwave absorption properties of FSS-impacted composites as a broadband microwave absorber. *Adv. Compos. Mater.* **26**(2), 99–113 (2017). <https://doi.org/10.1080/09243046.2016.1232009>
23. X. Chen, Z. Wu, Z. Zhang, Y. Zou, Ultra-broadband and wide-angle absorption based on 3D-printed pyramid. *Opt. Laser Technol.* **124**, 105972 (2020) (no. November 2019). <https://doi.org/10.1016/j.optlastec.2019.105972>

Chapter 3

Analytical Simulation of Spatial Distribution Lambertian Point-Source in LED Chip's Epitaxial Layers



Ahmad Noor Abdul Hamid, Ahmad Fakhurrazi Ahmad Noorden,
and Muhammad Zamzuri Abdul Kadir

Abstract In this paper, the theoretical simulation of spatial distribution Lambertian point source is performed by considering the refractive index of different material used. The refractive index contributes in changing the spatial light intensity produced at air layer of LED chip. The system required sandwich design of LED chips epitaxial layer which consist of P and N type material. At the top of the chip, Indium tin oxide (ITO) as current spreading layer has been placed. The formulation of Snell's law, total internal reflection and Lambertian emission function contributes to analyze the effect of different refractive index toward spatial light intensity produced at air layer. The change of the total internal reflection and spatial light intensity is analyzed which respect to the change of refractive index.

3.1 Introduction

In early twentieth century, a phenomenon named electroluminescence has been accounted due to light emission from a solid-state material, caused by an electrical power source. Started at room temperature, the electroluminescence can occur, it is on a very basic level unique in relation to incandescence (or heat glow), or, in other words electromagnetic radiation transmitted by a material warmed to high temperatures, ordinarily more than 750° [1]. Light-emitting diodes (LEDs) are semiconductor devices which can create ultraviolet, visible or infrared radiation. These essential properties make it become absolute source for multiple application such as light indicator, displays to solid-state lighting and optical communication system [2]. The LED is the recent light source with many advantages such as long lifetime, environment friendly, power saving and no pollution compare to others light source [3–5].

A. N. Abdul Hamid · A. F. Ahmad Noorden (✉) · M. Z. Abdul Kadir
Centre for Advanced Optoelectronics Research (CAPTOR), Kulliyah of Science, International
Islamic University Malaysia, Pahang, Malaysia
e-mail: fakhurrazi@iium.edu.my

M. Z. Abdul Kadir
e-mail: zamzurikadir@iium.edu.my

The LED is designed with modern technology and the advancement in fabrication process, device design and assembly technology that help to rise its performance and reduce its cost [6]. Recently many researches have been performed for optimizing the output light distribution and obtaining high efficiency of the lighting device [7–10]. However, it has been found that there are lack of analytical analysis or simulation which contribute to the model for the epitaxial layer of the LED chips. There also low extraction efficiency caused by the total internal reflection between active region and other layers.

In this paper, analytical simulation of spatial distribution Lambertian point-source was performed in order to observe the spatial light intensity distribution produced by LED chip's epitaxial layers. The Lambertian function is used to calculate the light intensity distribution. At first, a programming script is developed to execute the analytical simulation of spatial light intensity distribution across various angle of the Lambertian source. Besides, different materials were used for the chip's epitaxial layers in this simulation in order to compare the light intensity distribution for every material used. During this research, a mathematical based software is used in order to execute the modelling of the light distribution. Basis formulation including Snell's law, total internal reflection and Lambertian emission pattern were used. The square chip design is used in the simulation with 1 mm length and 1 mm width.

3.2 Analytical Formulation of Surface Emission Model for Led Chip

Based on previous research, there are many formulae that was used to calculate efficiency of the LED performance. A high efficiency is defined as a minimal amount of wasted time, capacity and materials. In this paper, the analytical formulation of surface emission model for LED chip involves the basis formulations including Snell's law, total internal reflection law, and Lambertian emission pattern.

3.3 Total Internal Reflection in Led Chip

The phenomenon known as the total internal reflection (TIR) is occur when the incident ray enters two-different medium with different refractive indexes. Then, when the angle of incidence ray reaches a critical value, the refracted ray's lies along the boundary which makes the angle of refraction turn to 90°. In short, this angle also known as critical angle. Using a mathematical expression, total internal reflection has been derived from the Snell's law [1]:

$$n_s \sin \Phi_i = n_{air} \sin \Phi_r \quad (3.1)$$

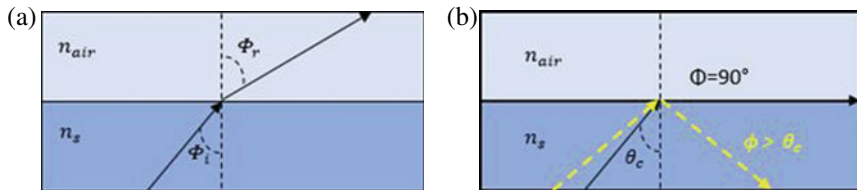


Fig. 3.1 Schematic diagram of **a** Light rays pass through two different media. **b** Total internal reflection phenomena

In this Eq. (3.1), n_s and n_{air} represent refractive index of semiconductor and refractive index of air respectively and the Φ_i and Φ_r represent the angle of refraction. This law explains what happen to the light source when it is going through the different media with different refractive index.

Figure 3.1a illustrated the light ray travelling from one media to another media with different refractive index. Normally, the n_s have higher refractive index compared to n_{air} . This different refractive index will affect the rate of transmission angle of the incident ray. The higher the different between the refractive index of the media, the higher the deflection of the transmission rays. Since the angle of refraction become 90° sine of $\Phi = 1$. The critical angle then given as;

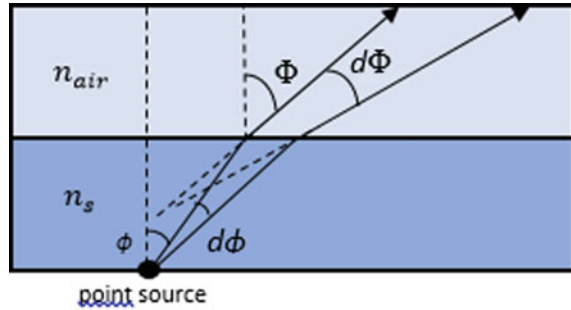
$$\theta_c = \sin^{-1} \frac{n_{air}}{n_s} \quad (3.2)$$

Equation (3.2) is the result when Snell's law was rearranged. In this equation, n_s and n_{air} represent refractive index of semiconductor and refractive index of air respectively and the θ_c and Φ represent the critical angle and transmission angle respectively. This law explains what happen to the light source when the incident angle reaches critical angle and what happen when incident angle, ϕ bigger than critical angle θ_c . When the incident angle reaches the critical angle, θ_c the transmission angle, Φ become 90° or become flat at semiconductor-air surface, that mean, at critical angle there is no transmission of light in second medium. When the incident angle ϕ bigger than critical angle θ_c , the transmission angle will totally be reflected into the same media. This phenomenon also known as total internal reflection (TIR) as shown in Fig. 3.1b.

3.4 Lambertian Emission Function for Planar Led Chip

Light particle also known as photon is emitted when recombination of holes and electron take place in active region of the LED chip. Due to the current only allowed to flow in forward bias, the electron and holes was forced towards active region and caused recombination to occur there. The photon produced then form the radiation emission pattern outside the LED chip [11].

Fig. 3.2 Schematic diagram of light rays emitted from the source passed through two different media



The Lambertian radiation pattern is produced when the LED chip layer of high index with planar surface [11]. In this research, the position of point source was considered based on the design and size of the LED chip with 1 mm length 1 mm width and 1 mm long. The position of the point source affect the Lambertian radiation pattern produced based on the formulation given.

$$I_{air} = \frac{P_{source}}{4\pi r^2} \frac{n_{air}^2}{n_s^2} \cos\Phi \quad (3.3)$$

The output intensity of the Lambertian radiation pattern was obtained using Eq. (3.3), where I_{air} represent a light intensity in air, P_{source} is the power of the light source provided, n_s and n_{air} represent refractive index of semiconductor and refractive index of air respectively, ϕ is the light emitted angle in the semiconductor and Φ is the transmitted angle in the air or second medium. Light also emitted into $d\phi$ in the semiconductor and transmitted into the air at $d\Phi$. The light intensity from the point light source follows a cosine dependence on the angle of light emitted, Φ . By following this law, the light intensity is maximum when the emission is normal to the surface or at angle $\Phi = 0^\circ$. The light intensity is always different at every angle of emission of light [11]. Figure 3.2 show how the light travel from one media to another.

3.5 Relation Between Position of Point Source and Intensity Produced

A light beam that was emitted from active region was divided into three categories, that is, some of the light beam was emitted through the chip, some are transmitted by refraction and some are reflected back into the chip.

In this paper, the LED chip layer was considered with two layers that are gallium nitride (GaN) and Indium Tin Oxide (ITO) as current spreading and air layer. The light intensity was investigated at both ITO and air layer. The point source was stated at 0, 0.5 and 1.0 mm on the chip layer in order to see the special light intensity

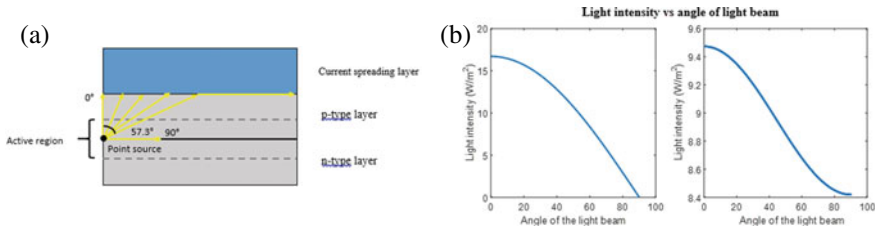


Fig. 3.3 a Schematic diagram of the position point source depends on angle of incident ray with critical angle and b light intensity against angle of the light beam produced at ITO layer (left) and air layer (right)

produced. As point source stated at different position, different spatial light intensity distribution has been produced in each layer based on the Eqs. 3.2 and 3.3.

However, based on Fig. 3.3a, the angle of refracted ray become constant (90°) after the incident ray are at 57.3°. Snell’s law states that when the angle of incidence ray is equal to the critical angle, the angle of refraction is at 90°. This shows that 57.3° is the critical angle at GaN layer in LED chip package. Although, it can be said that the critical angle is depending on the different refractive index of every different material used. Based on Fig. 3.3b, it shows that when the angle of refracted rays increases, then the light intensity at the epitaxial layer were decreases. Using Lambertian radiation pattern of Eq. 3.3, this equation follows a cosine dependence on the refracted angle θ . Besides, the pattern of light intensity produced is same at different position of point-like source at the epitaxial layer of LED chip. This is because, when the angle of refracted ray bends away from the normal line, then the intensity becomes lower. The angle of refracted ray for each position of point source were also be examined from 0° to 90° for every layer.

3.6 Material Used for Led Chip Layers

See Table 3.1.

Table 3.1 The table shows the list of material used for LED chip

Semiconductor material	Reflective index	References
Gallium (III) nitride (GaN)	2.500	[12]
Indium gallium nitride (InGaN)	2.590	[1]
Gallium (III) phosphide (GaP)	3.000	[11]
Aluminium gallium arsenide (AlGaAs)	3.250	[11]
Aluminium gallium indium phosphide (AlGaInP)	3.450	[1]
Aluminium gallium nitride (AlGaN)	3.810	[1]

3.7 Spatial Light Intensity Distribution by Different Refractive Index Material

Spatial light intensity distribution scientifically means how the pattern and how high the distribution intensity been produces across a spatial displacement. It also defines as how far the area of the light intensity being distributed inside and outside of the LED chip. The distribution pattern of spatial light intensity affects the area and the brightness of the light intensity. When the distribution pattern is large, the area of light produces also large and lead to the high brightness of light intensity. In this research, the simulation has been designed in order to investigate the spatial light intensity produced by the LED chip according to the angle of light refracted, the position of the point like source and different material of the chip epitaxial layers.

Based on the Fig. 3.4a–f, the pattern of the graph produce is the same, but it has different value of the maximum light intensity due to the different material with different refractive index has been used for the LED chip epitaxial layer. In this research, all six materials have been tested for the epitaxial layer of the LED chip. These six materials are, GaN, InGaN, GaP, AlGaAs, AlGaInP and AlGaN with refractive index 2.50, 2.59, 3.00, 3.25, 3.45 and 3.81 respectively. For low refractive

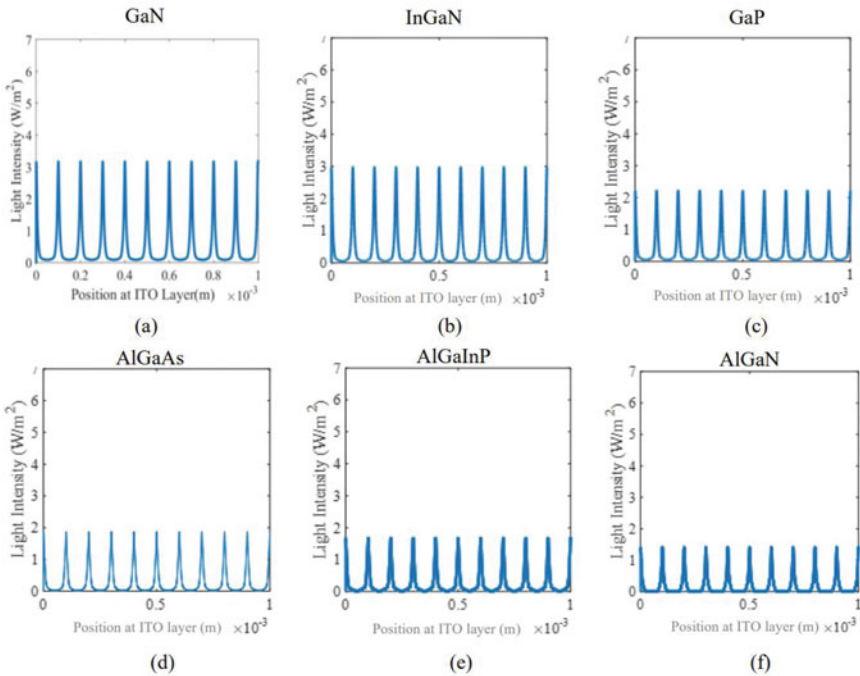


Fig. 3.4 Spatial distribution of light intensity at air layer from **a** GaN, **b** InGaN, **c** GaP, **d** AlGaAs, **e** AlGaInP, and **f** AlGaN material

index used for the LED chip epitaxial layer such as GaN, the light intensity produce at air layer is ~ 3.2 and for high refractive index such as AlGaIn is used, the light intensity produce is ~ 1.40 . Same for the spatial light intensity produced by side chip ray tracing system, for low refractive index used for the LED chip epitaxial layer such as GaN, the light intensity produce at air layer is ~ 3.1 and for high refractive index such as AlGaIn is used, the light intensity produce is ~ 1.40 . For side chip ray tracing system, to produce light intensity, same analytical formulation had been used including Eqs. 3.1–3.3 [1]. From the result obtain, it can be concluded that, when the refractive index of the material used for the LED chip epitaxial layer is high, then the light intensity produce at both ITO layer and Air layer is low.

3.8 Conclusion

In this paper, the highlight of this simulation project is focused on how spatial light intensity distribution were estimated by positioning the Lambertian point-like sources. It is used to identify the best material for the epitaxial layers of the LED chip for obtaining the high efficiency lighting source. The analytical formulation of Snell's law, total internal reflection and spatial light intensity distribution were involved.

Moreover, with some research of previous work, it shows that the light extraction efficiency ratio of ITO with a refractive index of 2.1 is the highest among other materials for current spreading layer. Thus, ITO was chosen as a current spreading layer in LED chip package in the simulation. Moreover, the highest value of spatial light intensity distribution is obtained from GaN based chip, in which for ITO layer is $\sim 14.0 \text{ W/m}^2$ and for air layer is $\sim 3.2 \text{ W/m}^2$. The construction of spatial light intensity distribution is depending on the position of point-like source and the angle of the light rays.

Acknowledgements This work was supported by the CAPTOR and Department of Physics, Kulliyah of Science, International Islamic University Malaysia, and the Ministry of Education (Malaysia) through Fundamental Research Grant Scheme (Project No.: FRGS19-033-0641) (Reference No: FRGS/1/2018/TK07/UIAM/02/1)

References

1. E.F. Schubert, *Light-emitting diodes* (Cambridge University Press, Cambridge, 2006)
2. I. Moreno, in *Spatial Distribution of LED Radiation*. International Optical Design Conference 2006. International Society for Optics and Photonics (2006)
3. D. Li et al., in *Numerical Simulation on Heat Pipe for High Power LED Multi-chip Module Packaging*. Electronic Packaging Technology & High Density Packaging, 2009. ICEPT-HDP'09. International Conference on. IEEE (2009)

4. E. Faraci et al., in *High Efficiency and Power Density GaN-Based LED Driver*. 2016 IEEE Applied Power Electronics Conference and Exposition (APEC). IEEE (2016)
5. B.U. Ye et al., Enhancing light emission of nanostructured vertical light-emitting diodes by minimizing total internal reflection. *Adv. Func. Mater.* **22**(3), 632–639 (2012)
6. M. Nishikawa et al., in *An LED Drive Circuit with Constant-Output-Current Control and Constant-Luminance Control*. Telecommunications Energy Conference, 2006. INTELEC'06. 28th Annual International. IEEE (2006)
7. J. Jiang et al., Optical design of a freeform TIR lens for LED streetlight. *Optik-Int. J. Light Electron. Opt.* **121**(19), 1761–1765 (2010)
8. I. Moreno, C.-C. Sun, Modeling the radiation pattern of LEDs. *Opt. Express* **16**(3), 1808–1819 (2008)
9. M.A. Crowder, C. Zhan, P.J. Schuele, *Light emitting diode (LED) using three-dimensional gallium nitride (GaN) pillar structures with planar surfaces*. Google Patents (2014)
10. C.-C. Sun et al., Precise optical modeling for LED lighting verified by cross correlation in the midfield region. *Opt. Lett.* **31**(14), 2193–2195 (2006)
11. E.F. Schubert, *Light-Emitting Diodes*. E. Fred Schubert (2018)
12. X. Jin et al., Enhancing GaN led efficiency through nano-gratings and standing wave analysis. *Nanomaterials* **8**(12), 1045 (2018)

Chapter 4

Free Spectral Range, and Coupling Coefficient Analysis of Integrated Fano Resonance in All-Pass Micro-resonator



Mohammad Amirul Hairol Aman, Ahmad Fakhurrrazi Ahmad Noorden, Faris Azim Ahmad Fajri, Ahmad Noor Abdul Hamid, and Suzairi Daud

Abstract We demonstrate the simulation for all-pass microring resonators with integrated Fano resonance to obtain high Free Spectral Range (FSR). The simulation mainly performed to generate high FSR by varying 2 parameters that affect the system which are radius of the ring and distance between end-facet. Furthermore, the analysis is proceeded by a variation of coupling coefficient. The impact to the output spectrum due to variations of the parameters was observed and analyzed, focusing on the FSR. We also investigate the effect of coupling coefficient variation and its contribution to the optical performance of the microring. The highest FSR obtained from this simulation was 273.752 nm.

4.1 Introduction

Micro-resonator in an optical device consists of at least a ring and a straight bus waveguide attach to it [1, 2]. With evanescent field as medium in the micro-resonator, the sensitivity can be further increased and immune to electromagnetic disturbance [3, 4]. Aside from sensitivity, the limit of detection (LOD) also plays a major role in the sensing field. The Larger LOD is desirable as it has a wider range of detection [5]. LOD is associated with free spectral range (FSR) in a way where high LOD can be obtained with wide FSR [5], where FSR is the distance between two resonance peaks [6], and it varies as the refractive index changes [7]. Therefore, FSR can be regarded as one of the key points in the sensing field. Broader FSR can be accomplished through Fano resonance compare with normal microring [8].

M. A. Hairol Aman · A. F. Ahmad Noorden (✉) · F. A. Ahmad Fajri · A. N. Abdul Hamid
Centre for Advanced Optoelectronics Research (CAPTOR), Kulliyah of Science, International Islamic University Malaysia, Pahang, Malaysia
e-mail: fakhurrrazi@iium.edu.my

S. Daud
Laser Center, Ibnu Sina Institute for Scientific and Industrial Research, Universiti Teknologi Malaysia, Johor Bahru, Johor, Malaysia
e-mail: suzairidaud@utm.my

Fano resonance was applied in the sensing field as the ultrasensitive biosensor [9], biochemical sensor [10] and temperature sensor [11]. Fano resonance spectrum can be obtained by combining microring with interferometers two-beam interferometer [12] and Mach–Zehnder interferometer (MZI) [13]. Fano resonance also can be generated by altering the configuration of the microring or adjusting the pathway of the resonance such as additional air-holes [14] or reflector [15] in the waveguide and feedback coupled-waveguide [16]. Fano resonance was generated by merging the Fabry–Perot resonance with circulation resonance; thus asymmetric Fano-like line-shape, is obtained [2, 15, 17]. Fano resonance can be tuned by manipulating the phase difference of circulation resonance and Fabry–Perot resonance [12]. However, Fano resonance has low FSR [15] and this occurrence greatly reduces the LOD.

In this paper, simulation of the integration of the Fano resonance into the convectional microring to achieve the optimized FSR. An all-pass microring resonator with end-facet in the bus waveguide was used in the simulation system. The end-facet is necessary to obtain Fabry–Perot resonance and by merging with circulation resonance from the ring with suitable phase difference. The Fano resonance is successfully achieved by the simulated microring resonator system. The output spectrum was analyzed based on FSR and it was compared with convectional microring. The effect of varying coupling coefficient is analyzed toward understand the dynamic behavior of the output spectrum. The simulation was performed by using a mathematical based programming software where the wavelength range was set from 1400 to 1750 nm and the distance between end-facet from 5 to 35 μm .

4.2 Model & Theory

The all-pass configuration of microring resonator is used in the simulated system. For obtaining Fano resonance, convectional all-pass microring is not sufficient, thus an additional feature is applied to the system where end-facet is added to the bus waveguide. The configuration is shown in Fig. 4.1 with labelled propagating electric

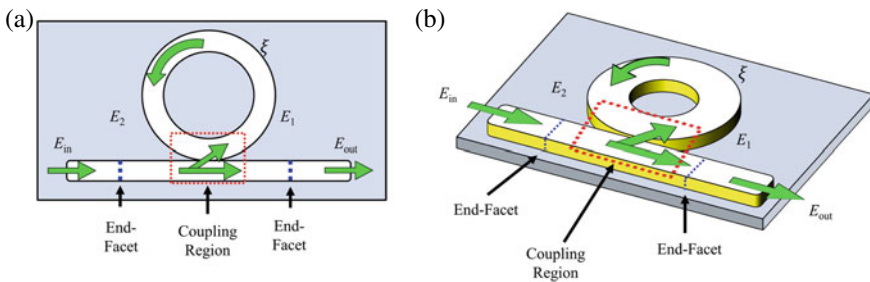


Fig. 4.1 a 2-D model and b 3-D model of all-pass microring resonator with additional end-facet in the waveguide. The green arrow in (a) and (b) represents the pathway of light in the microring while the red dotted box represents the coupling region

field. Light enters through the input port, E_{in} and propagates in the waveguide. At the coupling region, a fraction of the light was coupled into the ring at port E_1 , and another fraction keep propagates in the waveguide. The propagating light in the waveguide, partially reflected back due to the presence of end-facet and partially went through output port E_{out} . On the other hand, light in the ring propagates toward the coupling region at port E_2 and coupled back into the waveguide. Then, the resonance from ring will merge with the resonance in the waveguide, thus Fano-like resonance can be obtained. The end-facet was necessary for generating the Fabry–Perot resonance. The combination of circulation resonance and Fabry–Perot resonance will create a Fano resonance as the output spectrum.

To simulate the process, there is some mathematical formulation need to be considered. The transfer matrix for the whole system [17] can be expressed as

$$\begin{bmatrix} E_{out} \\ E_1 \end{bmatrix} = \frac{1}{i\sqrt{1-r_h^2}} \begin{bmatrix} -1 & -r_h \\ r_h & 1 \end{bmatrix} \cdot \begin{bmatrix} e^{i2\pi n_{eff}L/\lambda} & 0 \\ 0 & e^{-i2\pi n_{eff}L/\lambda} \end{bmatrix} \cdot \begin{bmatrix} t_R & 0 \\ 0 & 1 \end{bmatrix} \\ \cdot \begin{bmatrix} e^{i2\pi n_{eff}L/\lambda} & 0 \\ 0 & e^{-i2\pi n_{eff}L/\lambda} \end{bmatrix} \cdot \frac{1}{i\sqrt{1-r_h^2}} \begin{bmatrix} -1 & -r_h \\ r_h & 1 \end{bmatrix} \cdot \begin{bmatrix} E_{in} \\ E_2 \end{bmatrix} \quad (4.1)$$

E_{in} is the input port, E_{out} is the output port, E_1 is the input port to the ring and E_2 is the output port of the ring. The term n_{eff} is the effective refractive index where it considers both refractive indexes of core and cladding. L represents the distance between end-facet in the waveguide and lastly, t_R is the Optical Transfer Function (OTF) of all-pass microring resonator. The amplitude reflection coefficient [15] can be expressed as

$$r_h = \frac{n_{eff} - 1}{1 - n_{eff}}. \quad (4.2)$$

The OTF of all-pass microring resonator [1, 2] was given as

$$\frac{E_{out}}{E_{in}} = \frac{c - ae^{i2\pi nL_R/\lambda}}{1 - cae^{i2\pi nL_R/\lambda}}. \quad (4.3)$$

a in Eq. (4.3) is single-pass amplitude and $a = \exp(-\alpha L_R)$ [1, 2, 7]. α is attenuation coefficient and L_R is the perimeter of all-pass microring resonator. c is self-coupling coefficient which can be represented as $c = \sqrt{1 - \kappa}\sqrt{1 - \gamma}$ [7] where κ is coupling coefficient [7] and γ is propagation loss coefficient [7].

4.3 Result and Discussion

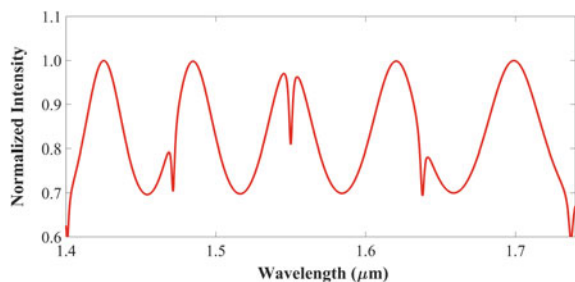
In this work, the material for the core and cladding of the proposed model were Silicon and Silicon Oxide, respectively. The system was simulated with wavelength range from 1400 to 1750 nm. The parameter of the system was varied where radius of the ring was simulated from 1 to 20 μm and distance between end-facet was simulated with range 5–35 μm . The coupling coefficient between the waveguide and the ring was 0.01. Figure 4.2 shows the spectrum of the microring with end-facet which have Fano-like line-shape.

This occurs due to integration between circulation resonance from the ring and Fabry–Perot resonance in the waveguide. Fabry–Perot resonance was formed by the partially reflecting wave which was reflected by the end-facet on both sides of the waveguide. Then, the reflected resonance combines with the circulation resonance from the ring, thus the spectrum with Fano-like shape as in Fig. 4.2 was achieved. This spectrum was similar to the Fano resonance where there was presence of asymmetrical shape in the spectrum with certain region has deep slope.

The system then was optimized to obtain the highest FSR within the wavelength domain by varying the radius of the ring from 1 to 20 μm and the distance between the end facet from 5 to 35 μm . Including the coupling coefficient which is set to 0.01 as a constants throughout the optimization process, the highest FSR achieved from this simulation was 273.752 nm. The FSR was obtained with 15.06 μm as the radius of the ring and the distance between end facet was determined as 9.2 μm .

The obtained FSR from the integration of Fano resonance into the microring was compared with circulation resonance from convectional microring to observe the impact of Fano resonance towards the output spectrum. With the same parameters as obtained by the spectrum of Fano resonance where the radius of the ring is 15.06 μm , distance between end-facet 9.2 μm and 0.01 as coupling coefficient, a new simulation was done for convectional microring. Figure 4.3b shows 2 output spectrums which are Fano resonance and circulation resonance with the absence of Fano resonance. FSR_1 and FSR_2 are the FSR of the resonance with and without Fano resonance, respectively. Based on Fig. 4.3a, FSR_1 is clearly more superior compared with FSR_2 where FSR_1 is 273.752 nm and FSR_2 is 83.595 nm. There is a difference of 69.46% between the 2 FSRs. Having high FSR also become an important factor need to be considered in an optical device. This is due to higher FSR contribute to a higher

Fig. 4.2 The output spectrum of integrated Fano resonance with circulation resonance



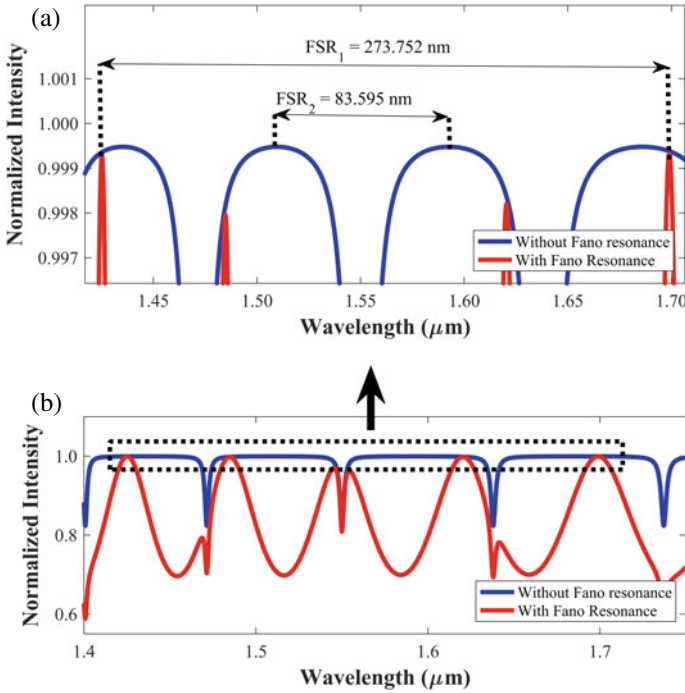
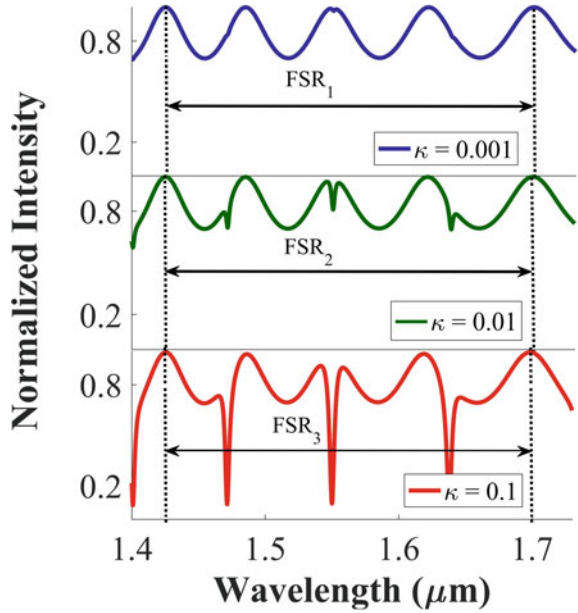


Fig. 4.3 **a** The magnified version for clear comparison of FSR between conventional all-pass microring and customized all-pass microring with end-facet in the waveguide. FSR_1 represents FSR of Fano resonance and FSR_2 represents the FSR of conventional circulation resonance. **b** The spectrum of conventional all-pass microring and customized all-pass microring with end-facet in the waveguide

limit of detection (LOD). Therefore, the integrated Fano-microring resonance has a higher limit of detection compared with the conventional microring.

The system was further study by varying the coupling coefficient, κ in order to observe the behaviour of the spectrum. From Fig. 4.4, the values of FSR with the variation of the coupling coefficients are negligible due to the difference among them is too small. The FSR for each spectrum was determined where FSR_1 is 273.84 nm, FSR_2 is 273.77 nm and FSR_3 is 273.11 nm. The slope of the resonance also become more deeper and steeper as coupling coefficient increases from 0.001 to 0.01 to 0.1. It is general knowledge that sensitivity can be measured by using the wavelength-shift technique and intensity-variation technique [18]. The former method can be used to calculate the sensitivity produced by Fano resonance as it has a deep slope in its spectrum. The steeper the slope of the resonance, the higher the sensitivity became [15]. The sensitivity of the spectrum for $\kappa = 0.01$ was measured to be 50.5 RIU^{-1} and for $\kappa = 0.1$, the sensitivity was measured to be 170 RIU^{-1} . Therefore, by utilizing the Fano resonance spectrum, increasing in coupling coefficient can contribute to higher sensitivity.

Fig. 4.4 The Spectrum of the resonance when the coupling coefficient, κ was varied from 0.001, 0.01 and 0.1



4.4 Conclusion

In conclusion, the integration of Fano resonance into the microring system through the additional feature of end-facet in the bus waveguide is a proven approach for obtaining an optical device with high FSR. The FSR of microring with integrated Fano resonance has wider FSR compare to FSR of convectional microring. The simulation was performed for the all-pass microring resonator configuration and the obtained value of FSR is 273.752 nm with parameter of the radius of the ring is 15.06 μm and the distance between end-facet is 9.2 μm . Variation of coupling coefficient affects the spectrum through the steepness of the slope. Increasing coupling coefficient make the steepness increase. Thus, the sensitivity of the optical device can be further enhanced by the integrated Fano resonance system. The system is tested for wavelength shift, and sensitivity is obtained as for $\kappa = 0.1$ is 170 RIU^{-1} which is higher when compared with $\kappa = 0.01$.

Acknowledgements This work was supported by CAPTOR and the Department of Physics, Kulliyah of Science, International Islamic University Malaysia, in term of facilities and financially by the Ministry of Education (Malaysia) through Fundamental Research Grant Scheme (Project No.: FRGS 19-033-0641) (References No.: FRGS/1/2018/TK07/UIAM/02/1).

References

1. W. Bogaerts et al., Silicon microring resonators. *Laser Photonics Rev.* **6**(1), 47–73 (2012)
2. J. Heebner, R. Grover, T. Ibrahim, *Optical Microresonator Theory* (Springer, 2008)
3. S.M. Chandani, N.A.J.I.P.T.L. Jaeger, Fiber-optic temperature sensor using evanescent fields in D fibers. *IEEE Photonics Technol. Lett.* **17**(12), 2706–2708 (2005)
4. G. Brambilla et al., Optical fiber nanowires and microwires: fabrication and applications. *Opt. Fiber Nanowires Microwires: Fab. Appl.* **1**(1), 107–161 (2009)
5. C. Wang et al., Large-dynamic-range dual-parameter sensor using broad FSR multimode photonic crystal nanobeam cavity. *IEEE Photonics J.* **10**(5), 1–14 (2018)
6. *Ring Resonators: Theory and Modeling, in Integrated Ring Resonators: The Compendium* (Springer Berlin Heidelberg, 2007), pp. 3–40
7. A.F.A. Noorden et al., Free spectral range analysis of double series microresonator system for all-optical corrosion sensor. *Opt. Eng.* **59**(1), 017106 (2020)
8. C. Qiu et al., Asymmetric Fano resonance in eye-like microring system. *Appl. Phys. Lett.* **101**(2), 021110 (2012)
9. K.-L. Lee et al., Ultrasensitive biosensors using enhanced Fano resonances in capped gold nanoslit arrays. *Sci. Rep.* **5**(1), 1–9 (2015)
10. X. Zhou et al., On-chip biological and chemical sensing with reversed Fano lineshape enabled by embedded microring resonators. *IEEE J. Sel. Top. Quantum Electron.* **20**(3), 35–44 (2013)
11. C. Qiu et al., in *A Temperature Sensor Based on Silicon Eye-Like Microring with Sharp Asymmetric Fano Resonance*. The 9th International Conference on Group IV Photonics (GFP). IEEE (2012)
12. T. Hu et al., Tunable Fano resonances based on two-beam interference in microring resonator. *Appl. Phys. Lett.* **102**(1), 011112 (2013)
13. L. Zhou, A.W.J.O.I. Poon, Fano resonance-based electrically reconfigurable add-drop filters in silicon microring resonator-coupled Mach-Zehnder interferometers. *Opt. Lett.* **32**(7), 781–783 (2007)
14. L. Gu et al., Fano resonance lineshapes in a waveguide-microring structure enabled by an air-hole. *APL Photonics* **5**(1), 016108 (2020)
15. H. Yi, D. Citrin, Z.J.O.E. Zhou, Highly sensitive silicon microring sensor with sharp asymmetrical resonance. *Opt. Express* **18**(3), 2967–2972 (2010)
16. G. Zhao et al., Tunable Fano resonances based on microring resonator with feedback coupled waveguide. *Opt. Express* **24**(18), 20187–20195 (2016)
17. L. Gu et al., A compact structure for realizing Lorentzian, Fano, and electromagnetically induced transparency resonance lineshapes in a microring resonator. *Nanophotonics* **8**(5), 841–848 (2019)
18. C.-Y. Chao, L.J.J.o.I.t. Guo, Design and optimization of microring resonators in biochemical sensing applications. *J. Lightwave Technol.* **24**(3), 1395–1402 (2006)

Chapter 5

Doping Effect on Bandgap Energy and Luminescence Spectrum for AlN-Based Semiconductor



Faris Azim Ahmad Fajri, Mohammad Amirul Hairol Aman, Ahmad Fakhurrrazi Ahmad Noorden, Ahmad Noor Abdul Hamid, and Azni Abdul Aziz

Abstract This ongoing work reports the heavy doping effect on the Aluminum Nitride (AlN) semiconductor (SC) material, illustrated via its energy-band structure (EBS). The research correlates the bandgap energy (BE) and depletion region (DR), which are then applied to the estimation of light-emitting diode (LED) luminescence spectrum (LS). The measurements are compared with different dopant concentrations ($1 \times 10^{18} \text{ cm}^{-3}$ – $1 \times 10^{21} \text{ cm}^{-3}$). Having the Gallium Arsenide (GaAs) properties as the controlled variable, the EBS is validated with literature findings. The measured band gap energy of AlN shifts from 6.2435 to 6.2326 eV. It decreased as the dopant concentration increased. However, the active spatial regions, reduced from $(1.0250 \times 10^{-1} \text{ to } 4.5000 \times 10^{-3} \mu\text{m}) \times 1 \mu\text{m}^2$. The findings are compared with the output LS of LED using the same SC material. The changes in BE and DR are consistent with the LS peak intensity wavelength and relative intensity to all the chosen doping concentrations. Though acquiring this consistency, an extensive discussion with collaboration in material science studies will further strengthen the understanding regarding these behaviours.

5.1 Introduction

Understanding doping in semiconductor (SC) materials is crucially valuable. The process is essential, if not a requirement in the fabrication of current semiconductor materials, as it contributes to the SC conduction mechanisms. Semiconductor group III-V such as Gallium Nitride (GaN) and Aluminum Nitride (AlN), are widely studied

F. A. Ahmad Fajri · M. A. Hairol Aman · A. F. Ahmad Noorden (✉) · A. N. Abdul Hamid · A. Abdul Aziz

Centre for Advanced Optoelectronics Research (CAPTOR), Kulliyah of Science, International Islamic University Malaysia, Pahang, Malaysia
e-mail: fakhurrrazi@iium.edu.my

A. Abdul Aziz
e-mail: azniabdulaziz@iium.edu.my

in high energy mobility transistor (HEMT), light-emitting diodes (LED), laser diodes, and solar cells [1–5]. We have included some of their unique properties in Sect. 5.2, where generally, the former operates at higher bias potential, including higher power density than the latter [6].

Furthermore, to aid in understanding the mechanism of these devices, specific numerical systems have been established and advanced [7–9]. Most of the dynamical mechanisms of the carriers are connectively explained via its energy band structure (EBS) or deductively conclude by its efficiency calculations [10, 11]. Recently, material variables contributing to the understanding such as the heterostructure configuration, thermoelectric properties, piezo-polarization effects, the effective mass of carriers, etc., are being utilized in the semiconductor’s mechanism [12–15].

However, the explanation involving semiclassical formulation in the doping process effects to the overview of EBS is still present with a large gap. Recently, K. P. Misra et al. and A.M. Robert et al. presented new interesting modellings to be considered for the bandgap energy (BE) comprehension [13, 16]. Yet, the mathematical derivation of BE and DR based on the EBS, along with the correlations of the dopants’ concentrations, are vague. Thus, in this paper, the focus is directed to the formulation of these properties in AlN semiconductors. Also, the correlations are strengthened by visualizing the effects on the LED luminescence spectrum (LS) for each material. Finally, the description of the numerical simulation is explained by introducing the modelling and approximations used to understand the main inquired variables, the BE, DR, and LS.

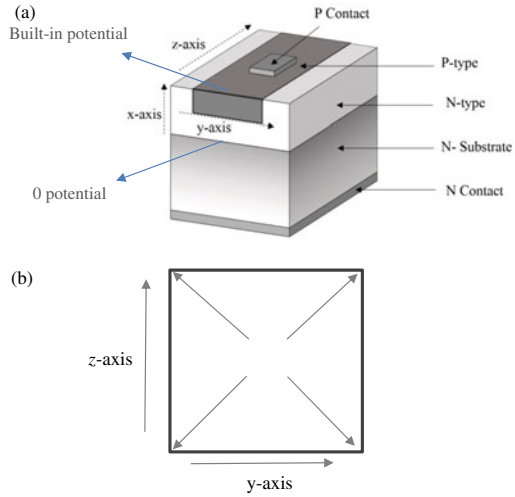
5.2 Simulation Model

5.2.1 Epitaxial Layer Modelling and Approximation

The mathematical modelling for EBS construction, drift–diffusion model (DDM) is based on three fundamental equations, which are the Poisson’s equation (description of the holes and electrons distribution in the semiconductor), current equation (represents the rate of current changes in the semiconductor), and drift–diffusion equation (formulation of how the carriers flow in the epi-layer) [7, 17, 18]. The equations are self-consistently solved using the finite difference method (FDM) and iteration by the Newton–Raphson method.

Several approximations have been made in this ongoing work: The semiconductor modelling is solved by applying homogeneity across the semiconductor; Altering the dopants concentration affects the effective mass of electrons only; The boundary condition of the forward-biased at the ends of the epitaxial layers. Thus, based on the configuration in Fig. 5.1, the modelling is linearised so that the carriers evolve homogeneously at each specific nodal point on the x-axis through the n-p region.

Fig. 5.1 Schematic diagram of the N-P region where **a** 3D view of the region and **b** is the area of a single nodal point, where the surface area is normal to the x-axis



5.2.2 EBS Construction

The solutions by solving the DDM are obtained to construct the EBS. Solutions such as the fermi potentials of AlN and its carriers' concentrations are applied to valence, E_v (Eq. 5.1) and conduction bands, E_c (Eq. 5.2). The spatial dimension is set as the domain, and the bands are formulated as the following.

$$E_v(x) = k_B T \ln\left(\frac{p(x)}{N_v}\right) - q\varphi_p(x) \quad (5.1)$$

$$E_c(x) = -q\varphi_n(x) - k_B T \ln\left(\frac{n(x)}{N_c}\right) \quad (5.2)$$

where p and n are the holes and electrons concentration, φ_p and φ_n are the quasi-Fermi levels, k_B is the Boltzmann's constant, q is the single carrier charge, T is the absolute temperature, and N_v and N_c are the effective density of states (DOS). N_v and N_c are dependants on the effective mass of holes and electrons, respectively, including the temperature.

The DR is derived from EBS; visually, in the region where the bands are changing. This is because DR is the space where the recombination of holes and electrons occurs. From quantum mechanics perspectives, the carriers shift between the available states while recombine and regenerate, producing photons during the diffusion across the epitaxial layer. Mathematically, it is measured by the following expressions

$$R = \int_{S_2}^{S_1} dV \quad (5.3)$$

$$\left| \frac{d}{dV}(E_c(V), E_v(V)) \right| \geq \delta, \text{ where } S_1 \text{ and } S_2 \text{ must satisfy} \quad (5.4)$$

$$|E_v((S_1, S_2) \pm \Delta S)| \geq \delta; |E_c((S_1, S_2) \pm \Delta S)| \geq \delta \quad (5.5)$$

δ is the minimum energy of recombination in a semiconductor, V is the volume of the epitaxial layer, S_1 and S_2 are the nodal points closest to the boundary of where the recombination occurs, and ΔS is the distance between neighbouring nodal points. Based on the linearisation of DDM, it led to a more straightforward way to determine the region via the following form

$$A \int_{x_2}^{x_1} dx \quad (5.6)$$

Based on the equivalent condition from Eqs. (5.4), and (5.5), the calculations depend on A as the surface area of the epitaxial layer, x_1 and x_2 are the nodal points closest to the end of the DR, respectively.

Applying these effects to the estimation of LS is possible with the available solutions obtained by solving the DDM. The product of the joint density of states of the carriers and the Boltzmann's distribution function gives LS, which is as the following [11, 17].

$$I(E) = \left(\frac{1}{2\pi} \left(\frac{2m_r^*}{\hbar^2} \right)^{3/2} \sqrt{E - E_g} \right) * (e^{-E/(k_B T)}) \quad (5.7)$$

where \hbar is a form of the Plank's constant, and m_r is the reduced mass of the carriers, holes, and electrons. For the spectrum estimation, the equation is only valid for positive values of E_g calculated from EBS. A negative value of calculated E_g , signifies that the SC material has behaved as the non-SC (conductor material). Recombination of carriers is unlikely to occur in conductor material.

5.2.3 Simulation

The simulation was conducted in the GNU Octave environment initialized by the material properties. The simulation calculates the BE and DR based on different dopant concentration values, N_A the acceptor concentration, and N_D the donor concentration. The simulation parameters referred from [11, 18–20] for AlN are recorded in Table 5.1.

Validation. The DDM simulation system is validated by comparing the simulated result with literature findings for GaAs materials. The comparison is performed by computing the ratio of this research data to the literature data [7]. Both simulations

Table 5.1 Simulation parameters for AlN at T = 300 K

Properties	AlN	Units
Intrinsic carrier concentration, n_i	9.4×10^{-34}	cm^{-3}
Bandgap Energy	6.23	eV
Effective DOS in conduction band, N_c	2.23×10^{18}	cm^{-3}
Effective DOS in valence band, N_v	4.62×10^{19}	cm^{-3}
Mass of electron at rest, m_0	9.12×10^{-31}	kg
Effective mass of electrons, m_e	$0.4 \times m_0$	kg
Effective mass of holes, m_h	$7.26 \times m_0$	kg
Electron mobility, μ_e	300	$\text{cm}^2/(\text{Vs})$
Hole mobility, μ_h	14	$\text{cm}^2/(\text{Vs})$
Non radiative time constant, τ_n	1×10^{-9}	s
Relative dielectric constant, ϵ	8.5	
<i>n</i> -type surface area ^a	1×1	μm^2
<i>p</i> -type surface area ^a	1×1	μm^2
Thickness	1	μm

DOS—Density of states

^aSurface area is normal to the *x*-axis

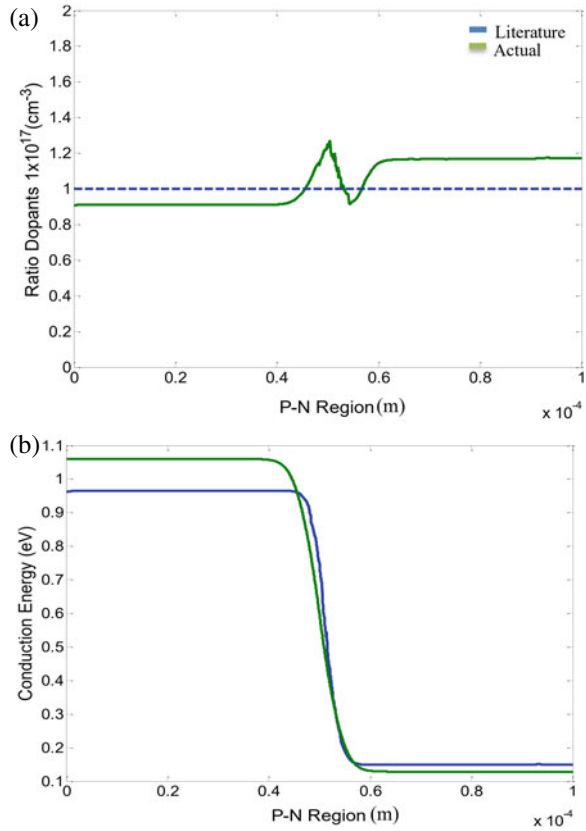
were operated based on the same parameters recorded in the same literature. The results in Figs. 5.2a and 5.3a are for the dopant concentrations of $1 \times 10^{17} \text{ cm}^{-3}$ and $1 \times 10^{18} \text{ cm}^{-3}$, respectively. Similarly, for the energy band, the comparison is in Figs. 5.2b and 5.3b. Referring to the comparisons in Figs. 5.2a and 5.3a, the differences between the literature data and simulated results occurred between 0.4 and 0.6 ($\times 10^{-4} \text{ cm}$) of the **P-N region** (opposite to our N-P configuration for AlN simulation). The blue line in the graphs reflects the reference line representing the value of the ratio equal to 1. The difference is still acceptable since it fluctuates only in a specific domain range, especially in Fig. 5.2a).

For energy band comparison as in Figs. 5.2b) and 5.3b), it is visible that the solutions are consistently following the trend of the literature data. It is also notable that there are differences between them at the point when the conduction energy changes. The differences occur in the EBS graphs as the simulated data is slightly shifted to the right by less than 0.1 μm . Due to the drastic changes on the vertical axis, the fluctuation of the ratio becomes drastic. There are minor errors in the active region which covers less than 0.2 μm of the epitaxial layer.

In addition, there are slight differences in the actual data for both doping concentrations (Figs. 5.2b and 5.3b), where the height of conduction bands of the existing data is visibly shorter for both P and N regions of the epitaxial layers. However, the differences are minutely stable, as shown in the ratio visualization.

The differences may occur due to three main reasons, which are (1) a smaller number of nodes chosen in our simulation, (2) the iteration flow in the Newton

Fig. 5.2 The validation graph of **a** the ratio of the literature data to the simulated data and **b** the energy band structure of $1 \times 10^{17} \text{ cm}^{-3}$ doping concentration



methods, and (3) the most probable factor is that we simulate a non-constant effective mass value while the referred data simulate constant values for the carriers. Nonetheless, by comparison, the system is in good agreement in terms of carriers' distribution in the 1D P-N region.

5.3 Energy Band Structure

The EBS are constructed in Fig. 5.4, according to the colour from black ($N_A, N_D = 1 \times 10^{18} \text{ cm}^{-3}$), blue ($N_A, N_D = 1 \times 10^{19} \text{ cm}^{-3}$), green ($N_A, N_D = 1 \times 10^{20} \text{ cm}^{-3}$), and red ($N_A, N_D = 1 \times 10^{21} \text{ cm}^{-3}$) in increasing order.

Based on the approximation where the homogeneity is conserved through each node along the x -axis, the increasing amount of dopants concentration reduces the recombination region between carriers. Thus, logically, a reduced space will lower the probability of the recombination process occurring. This claim is supported by the estimation of LS in the following subsections.

Fig. 5.3 The validation graph of **a** the ratio of the literature data to the simulated data and **b** the energy band structure of $1 \times 10^{18} \text{ cm}^{-3}$ doping concentration

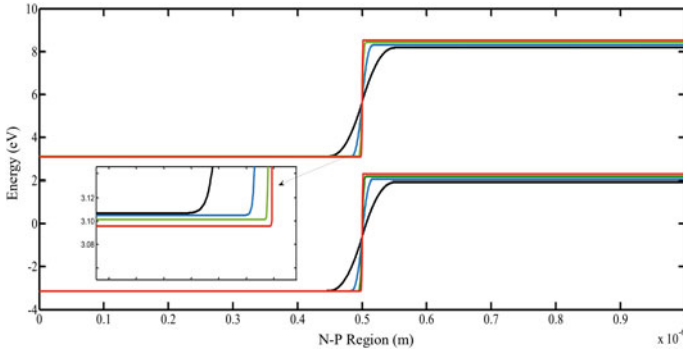
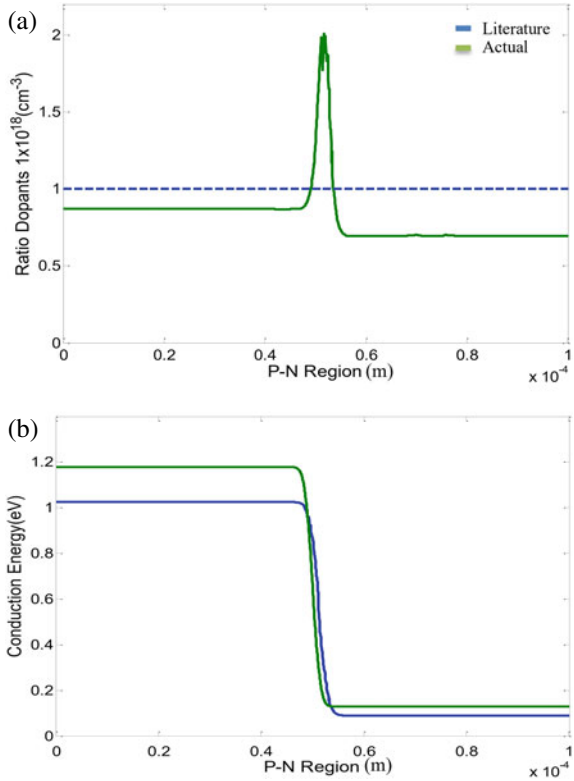


Fig. 5.4 The Energy Band Structure of AlN consists of E_c (upper group) and E_v (lower group). The bands, E_{c18} to E_{c21} and E_{v18} to E_{v21} are based on dopants concentration values

However, the differences of E_v from E_c , i.e. $(E_c - E_v)$, are observably constant in Fig. 5.4 since the gap between the same-coloured line seems to stay the same. The average gap for semiconductors is between ~ 0 eV (conductor) to >15 eV (insulator) [21], so this signifies that AlN can maintain its semiconductor properties. In other words, it does not behave like a non-SC though it is doped with a high dopant concentration. The overall view is analyzed in detail by calculation to determine whether the gap differs or remains the same. Nevertheless, as an overview, EBS is helpful to visualize the general semiconductor behaviour where each material forms its own unique EBS.

5.3.1 Bandgap Energy

The calculation of BEs is recorded in Fig. 5.5, and several chosen data extracted from it are tabulated in Table 5.2. The BE, E_g values of each dopant concentration are calculated based on the average differences. In Fig. 5.5, the BE values show drastic changes starting from the dopants concentration of $1 \times 10^{18} \text{ cm}^{-3}$, where

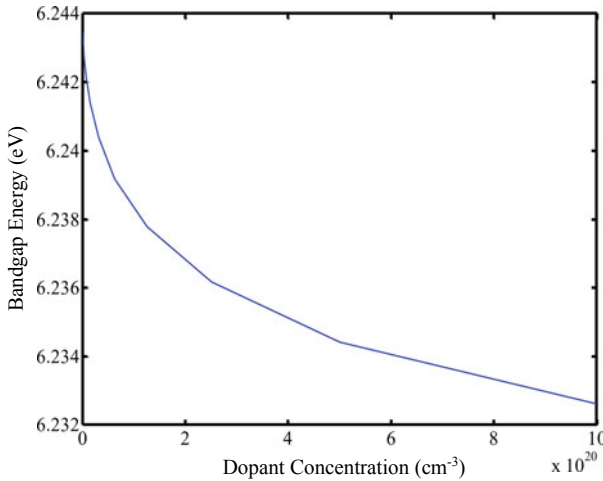


Fig. 5.5 The band energy against the dopants concentration of AlN

Table 5.2 Bandgap energy AlN results from dopants concentrations change

Dopants concentration (cm^{-3})	Bandgap energy (eV)
1.000×10^{18}	6.2435
1.000×10^{19}	6.2419
1.000×10^{20}	6.2383
1.000×10^{21}	6.2326

E_g decreases by 0.0016 eV from the initial value in Table 5.1. A blue shift of BE has occurred. The BE is then reduced more by 0.0036 and 0.0057 eV for the subsequent increments of dopants concentration. A further investigation with the additional knowledge of material physics experts will undoubtedly provide solid fundamental understandings for this study. Nevertheless, based on the graph trend, the shift will eventually be saturated as it is doped with a higher concentration [22, 23].

However, due to limited computational resources, $1 \times 10^{21} \text{ cm}^{-3}$ is the highest doping concentration our simulation system can solve. In quantum mechanics, this effect can be understood from the description of the Moss-Burstein effects explain the behaviours. Generally, it can be summarised from [21, 24, 25] that the BE shifts due to the occupied quantum states of the original valence/conduction band, and eventually, the current overflow occurs as the unnecessary carriers are still injected/doped.

5.3.2 Depletion Region

The DR is the space where recombination of carriers occurs; thus, it is the active region. Figure 5.6 shows the data for different dopant concentrations. The DR of AlN decreases rapidly as the dopant concentrations increase, approaching the saturation point at a value less than 100 nm. For clarity purposes, the chosen data are recorded in Table 5.3. The increment of doping concentration leads to a smaller active region. This condition probably occurred because of the same factor mentioned previously, where it lowers the BE of AlN by occupying the available quantum states.

Comparing the results recorded in the previous section, providing numerous amounts of carriers lowers the minimum energy required for the recombination process. But, unfortunately, it reduces the available space for the recombination process (active region), which lowers the probability of the process. Thus, this deteriorates the efficiency due to less photons production.

Fig. 5.6 The depletion region of AlN with respect to the dopant concentrations

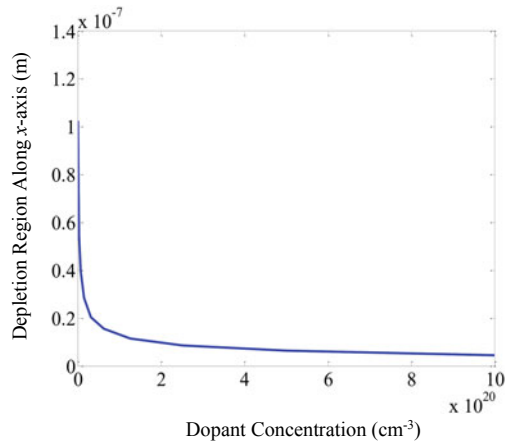


Table 5.3 DR of AlN by the increment of dopants concentrations

Dopants concentration (cm^{-3})	AlN (μm^3)
1.000×10^{18}	1.0250×10^{-1}
1.000×10^{19}	3.5500×10^{-2}
1.000×10^{20}	1.2500×10^{-2}
1.000×10^{21}	4.5000×10^{-3}

5.4 Luminescence Spectrum

The estimation of the LS is constructed in Fig. 5.7, based on the calculated E_g in Table 5.2. The colour of visible light changes with different wavelengths of light-wave, which is affected by the changes in BE. The detailed comparison is recorded in Table 5.4 in the nanometer unit range.

The doping of $1 \times 10^{18} \text{ cm}^{-3}$ shows the peak intensity wavelength, λ_{peak} emitted, is the longest and has the highest relative intensity, I_{rel} value than the other concentration of dopants. Both of these values decrease as the doping concentration is increased. On the contrary, by increasing the doping concentration, the rate of changes for

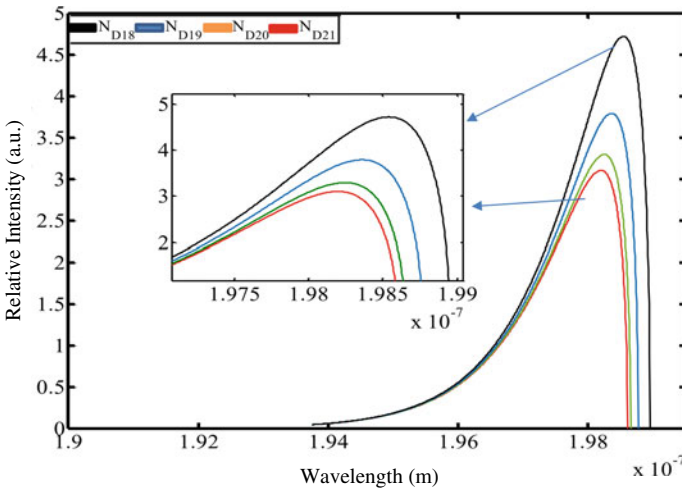


Fig. 5.7 The relative luminescence spectrum of AlN where N_D is the dopants concentration

Table 5.4 Peak intensity wavelength of AlN results from dopants concentrations change

Dopants concentration (cm^{-3})	AlN (nm)
1.000×10^{18}	198.1964
1.000×10^{19}	198.2457
1.000×10^{20}	198.3606
1.000×10^{21}	198.5407

λ_{peak} increases while the rate for I_{rel} decreases, noticeably in Fig. 5.7. A similar phenomenon is recorded by the blue-LED finding [26], where the ultraviolet-LED (UV-LED) of GaN emits the blue-coloured light by having the Indium as its dopant (that alters its bandgap energy).

Furthermore, to aid the understanding of the behavior of BE and DR properties, there are several literature that recorded the changes of properties in AlN based on the dopants concentration. As recorded in [27], the increment of Silicon (Si) concentration in AlN reduces the free electron concentration which is commonly referred to as the compensation knee. [24] reported that the transition energy level (ionization energy) in the Mg-doped AlN decreases when the doping concentration is increased. This supports our report as the bandgap energy is directly proportional to the ionization energy which has been thoroughly studied by [25].

Despite the lack of understanding of the factors affecting BE and DR results, we correlate the factor based on the output of the recombination process to show that it is consistent with these changes. Two aspects are considered:

Firstly, the measured photoluminescence of the LED depends on the E_g of the material. Thus, including the colour of light emitted that depends on the wavelength produced by the LED. The shift of wavelength is closely related to how the BE changes with the increment of dopant concentration. This is undoubtedly true as the LS justified mathematically via the 2nd term of Eq. (5.7), $(\sqrt{E - E_g})$ where the increasing BE will reduce the relative intensity of the SC material;

Secondly, the efficiency of LED mainly depends on the recombination process where the intensity of photon production can be. Therefore, the relative intensity values are connected to the behaviour of DR. The changes in DR, and maximum intensity values are equivalent to each other as the increment of doping concentration will decrease their values simultaneously. This indicates that fewer recombination and regeneration processes occur when the region is reduced, resulting in the less emitted light.

5.5 Conclusion

A numerical simulation that studies the correlations between BE, DR, and LS with the dopant concentration based on AlN materials properties, has been done. The results show that the BE decreases significantly when the dopants concentrations are increased. Similarly, DR decreases drastically to less than 100 nm when the dopant concentrations increase, where it can get to the point where the recombination process would unlikely occur. Thus, all of these clearly show that the doping of donors and acceptors in semiconductors significantly affects the crucial properties of the general semiconductor epitaxial layers.

Nevertheless, one must realize that using DDM; one can construct, calculate, or provide numerous information aside from EBS. In other words, DDM is not restricted or limited to EBS discussion only. Yet, our research can achieve fruitful findings by merely discussing the EBS. These preliminary findings open to more

extensive discussions based on the fundamentals of solid-state physics, including the application-wise to the various semiconductor devices.

Acknowledgements This work was supported by CAPTOR, Kulliyah of Science, International Islamic University Malaysia, and the Ministry of Education (Malaysia) through the Fundamental Research Grant Scheme (Project No.: FRGS 19-033-0641) (References No.: FRGS/1/2018/TK07/UIAM/02/1).

References

1. J. Shealy et al., in *Gallium Nitride (GaN) HEMT's: Progress and Potential for Commercial Applications*. IEEE, vol. 24th Annual (2002) pp. 243–246
2. S.P. Denbaars et al., Development of gallium-nitride-based light-emitting diodes (LEDs) and laser diodes for energy-efficient lighting and displays. *Acta Mater.* **61**(3), 945–951 (2013)
3. A.S.A. Fletcher, D. Nirmal, A survey of Gallium Nitride HEMT for RF and high power applications. *Superlattices Microstruct.* **109**, 519–537 (2017)
4. W. Metaferia, K.L. Schulte, J. Simon, S. Johnston, A.J. Ptak, Gallium arsenide solar cells grown at rates exceeding 300 $\mu\text{m h}^{-1}$ by hydride vapor phase epitaxy. *Nat. Commun.* **10**(1), 1–8 (2019)
5. A.R. Gobat, M.F. Lamorte, G.W. Mciver, Characteristics of high-conversion-efficiency gallium-arsenide solar cells. *IRE Trans. Mil. Electron. MIL* **6**(1), 20–27 (1962)
6. K.R.L.B.S.V.G.P.L.V. Robert, R. Seal II, J. Gambogi, Gallium. United States Geol. Surv. Bull. **1802-H**, H1–H35 (2017)
7. D. Vasileska, S.M. Goodnick, G. Klimeck, Computational electronics: semiclassical and quantum device modeling and simulation (2017)
8. S. Datta, *Quantum Transport: Atom to Transistor* (Cambridge University Press, 2005)
9. I. Pisarenko, E. Ryndin, Drift-diffusion simulation of high-speed optoelectronic devices. *Electronics* **8**(1) (2019)
10. A. Shaari, A.F. Ahmad Noorden, S.N. Mohamad, S. Daud, Geometrical analysis of light-emitting diode for enhancing extraction efficiency. *J. Sustain. Sci. Manag.* **15**(6), 68–84 (2020)
11. E.F. Schubert, *Light-Emitting Diodes* (2nd edn., Cambridge University Press, 2018)
12. Y. Wu, D.-B. Zhang, Z. Zhao, J. Pei, B.-P. Zhang, Enhanced thermoelectric properties of ZnO: C doping and band gap tuning. *J. Eur. Ceram. Soc.* (2020)
13. K.P. Misra, S. Jain, A. Agarwala, N. Halder, S. Chattopadhyay, Effective mass model supported band gap variation in cobalt-doped ZnO nanoparticles obtained by co-precipitation. *Semiconductors* **54**(3), 311–316 (2020)
14. X. Zhang et al., Enhanced performances of PVK/ZnO nanorods/graphene heterostructure UV photodetector via piezo-phototronic interface engineering. *Adv. Mater. Interfaces* **6**(23), 1–8 (2019)
15. V.P. Sirkeli, O. Yilmazoglu, F. Küppers, H.L. Hartnagel, Effect of p-NiO and n-ZnSe interlayers on the efficiency of p-GaN/n-ZnO light-emitting diode structures. *Semicond. Sci. Technol.* **30**(6), 065005 (2015)
16. R.A. Makin, K. York, S.M. Durbin, R.J. Reeves, Revisiting semiconductor band gaps through structural motifs: An Ising model perspective. *Phys. Rev. B* **102**(11), 115202 (2020)
17. A. Shaari, F.A. Ahmad Fajri, A.F. Ahmad Noorden, M.Z. Abdul Kadir, S. Daud, Optimizing the efficiency of gallium nitride-based light-emitting diodes from contact area of current spreading to electrode. *Microw. Opt. Technol. Lett.*, 1–5 (2020)
18. S. Sujecki, *Photonics Modelling and Design* (2014)
19. V. Bougrov, M.E. Levinshtein, S.L. Rumyantsev, A. Zubrilov, *Properties of Advanced Semiconductor Materials GaN, AlN, InN, BN, SiC, SiGe* (Wiley, New York, 2001), p. 207

20. J. Piprek, *Semiconductor Optoelectronic Devices* (Elsevier, 2013)
21. A. Dittmer, R. Izsák, F. Neese, D. Maganas, Accurate band gap predictions of semiconductors in the framework of the similarity transformed equation of motion coupled cluster theory. *Inorg. Chem.* **58**(14), 9303–9315 (2019)
22. M.F.N. Mohsen, Modification of Welge's method of shock front location in the Buckley-Leverett problem for nonzero initial condition. *Soc. Pet. Eng. J.* **25**(4), 521–523 (1985)
23. H.J. Welge, A simplified method for computing oil recovery by gas or water drive introduction and theoretical background. *Pet. Trans. AIME* **195**(13), 91–98 (1951)
24. Y. Peng et al., Tunable electronic structures of p -type Mg doping in AlN nanosheet. *Appl. Phys. Lett.* **116**(044306), 0–4 (2014)
25. C.A. Klein, Bandgap dependence and related features of radiation ionization energies in semiconductors bandgap dependence and related features of radiation ionization energies. *J. Appl. Phys.* **39**(4), 2029–2038 (2003)
26. J.Y. Tsao et al., Solid-state lighting: an integrated human factors, technology, and economic perspective. *Proc. IEEE* **98**(7), 1162–1179 (2010)
27. J.S. Harris et al., On compensation in Si-doped AlN on compensation in Si-doped AlN. *Appl. Phys. Lett.* **112**(152101), 15–20 (2018)

Chapter 6

Characterization of Electrical, Optical and Topological Properties of ITO Thin Films for Solar Cells Application



Nursyahirah Masdan and Ahmad Hadi Ali

Abstract Indium tin oxide (ITO) deposited on glass and Silicon (Si) substrate using direct current (DC) magnetron sputtering were analysed. This study characterized the electrical, optical and topological characteristics of ITO thin films at different deposition time. ITO thin films exhibit a low resistivity of $2.95 \times 10^{-3} \Omega \text{ cm}$ with conductivity of $(3.38 \times 10^2 \Omega \text{ cm})^{-1}$ at 1200 s. The optical properties of ITO thin films were measured using UV–Vis spectrophotometer. The transmittance results of ITO thin films from 400 to 1200 s deposition time were achieved above 96% corresponding to wavelength ranges from 600 to 1200 nm. The absorbance shows a good agreement with transmittance data. It is observed from atomic force microscopy analysis, increased deposition time of ITO, the surface roughness (R_q) decreased from 12.6 to 5.03 nm.

6.1 Introduction

Nowadays, transparent conductive oxides (TCOs) are commonly used in various technologies. TCOs have high transparency in a wide spectral range and low electrical resistivity [1]. Most common TCOs thin films are doped with tin, for instance, are ITO and fluorine-doped tin oxide (FTO) [1]. Selecting the right TCOs can help advance research in photovoltaics and electronic device technologies. [2]. ITO is the most appropriate TCOs candidate for solar cell application. ITO is being investigated for flat panel displays, light-emitting diodes, solar cells, and optical coatings [3]. Because of its excellent optical transmission, electrical conductivity, strong adhesion to substrates, and good stability in harsh environments, ITO is known as a

N. Masdan (✉) · A. H. Ali

Laser and Semiconductor Technology Research Group (LASERG), PDRS-COR, Department of Physics and Chemistry, Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, Johor, Malaysia

e-mail: hw200003@siswa.uthm.edu.my

A. H. Ali

e-mail: ahadi@uthm.edu.my

unique materia [4]. ITO is an n-type semiconductor with excellent transparency in the visible region and high conductivity, hence used as anti-reflecting and in solar cells [5]. The key advantages of ITO are its low resistivity, excellent transparency in the visible region, high reflection in the infrared region and high chemical stability. There are different substrates used to deposit ITO films for application in solar cells, for example, polyethylene terephthalate (PET), n/p-gallium nitride (GaN), n/p-Silicon (Si), and glass [6]. ITO films have good optical and electrical characteristics, allowing for greater energy generation [7].

Commonly thin metal films are widely used to make contact on semiconductor optoelectronic devices due to their excellent electrical conductivity. However, thin metal films are opaque and have high reflectance characteristics. It is an obvious limiting factor to be used as front contact of solar cells due to poor light transmission. TCOs such as ITO shows a suitable electrically conductive material with low absorption of light, wide bandgap, and has high transmission of light in the visible region and high reflection in the infrared region. Furthermore, ITO offered the best material available, with excellent conductivity and transparency, chemical stability, good surface morphology, and strong adherence to the substrate. Hence, these advantages make ITO thin films suitable for a front contact in solar cells compared to thin metal films.

The properties of ITO thin films depend on deposition methods. ITO films can be deposited using a variety of methods, including radio frequency (RF) and direct current (DC) magnetron sputtering [7, 8], electron beam physical evaporation [9], metal-organic chemical vapor deposition [10], and thermal evaporation [11]. In addition, because of the exceptional results obtained, the sputtering technique is mostly utilized to deposit ITO thin films. This technique produced high-quality thin films as well as deposited over a large surface area can be obtained [12]. This technique, either DC or RF sputtering, is the most used method for large-area thin film deposition because it is cost-effective and allows for the growth of films on large substrate areas [13]. This technique provides for control of its sputtering parameters, allow for the production of high-quality metal contacts at low vacuum pressure [14].

In this study, ITO thin films deposited by varying deposition time for application in solar cells. This study aimed to characterize the effect of different deposition times of ITO on its properties. The properties of ITO thin films were examined, including their electrical, optical, and topological properties. The influences of different deposition times of ITO thin films on these properties are reported.

6.2 Material and Method

Decon-90 solution mixed with deionized water was used to clean glass substrates. Then, the substrates were fully immersed in the solution while shaking the beaker for 5 min and cleaned using KIMTECH lens tissue. Then, the substrates rinsed thoroughly in the deionized water and blown dry with nitrogen (N₂) gas. They were

cleaned in boiled acetone at 55 °C for 5 min for Si substrates. Then, Si were rinsed in deionized water and blown dry with N₂ gas.

ITO deposited on substrates using the Q150RS DC magnetron sputtering method. ITO thin films were deposited onto silicon (Si) substrate and glass. The Si substrate with orientation of (100) and dimension 1 × 1.0 cm² while for glass had dimension 2.0 × 1.5 cm². The deposition chamber was pumped to base vacuum pressure of 3 [× 10] ^3 mbar. The electrical current of sputtering was set at 120 mA. Then, the sample holder was rotated to make sure the sputtered ITO films are uniformly deposited on the substrates. ITO thin films produced three sets of samples at different deposition times of 400, 800 and 1200 s.

Keithley-2400 power, Pro-4 Software, and a probing station used to investigate the electrical properties of ITO thin films. The electrical resistivity and conductivity of the samples were measured. An atomic force microscope (AFM) analysis was performed to examine the surface roughness of samples. The optical characteristics of the samples measured at room temperature using an Ultraviolet–Visible (UV–Vis) spectrophotometer (UV-3600 Plus) SHIMADZU. The optical spectra recorded in the 400–1200 nm wavelength ranges. Their optical transmittance, reflectance, and absorbance were obtained. Atomic Force Microscope (AFM) was used to analyse the topology of ITO thin films. The parameter for the analysis of the AFM measured was surface roughness (Rq).

6.3 Results and Discussion

Figure 6.1 shows the resistivity and conductivity of ITO thin films on Si at a various deposition time. Table 6.1 shows the resistivity and conductivity value of ITO thin films deposited on Si substrate. It is shown that deposition time of ITO thin films greatly influences their electrical properties. The results achieved show that ITO films resistivity decreased, the conductivity increased with deposition time. Equation 6.1 were used to calculate the conductivity of ITO thin films [15]:

$$\sigma = \frac{1}{\rho} \quad (6.1)$$

where σ is the conductivity, and ρ is the resistivity.

UV-Vis spectrophotometer used to characterize the transmittance and absorbance of ITO thin films on glass substrate. High transmittance achieved, especially in the visible light region, is essential properties for solar cells. From the data obtained, it can state that different deposition times of ITO significantly affect their transmittance. Figure 6.2 shows the transmittance and absorbance of ITO thin films against wavelength. As expected, ITO thin films transmittance increased with an increase in deposition time. At all wavelength bands, all samples produce greater than 96%

Fig. 6.1 The electrical resistivity and electrical conductivity of ITO thin films deposited on Si against deposition time

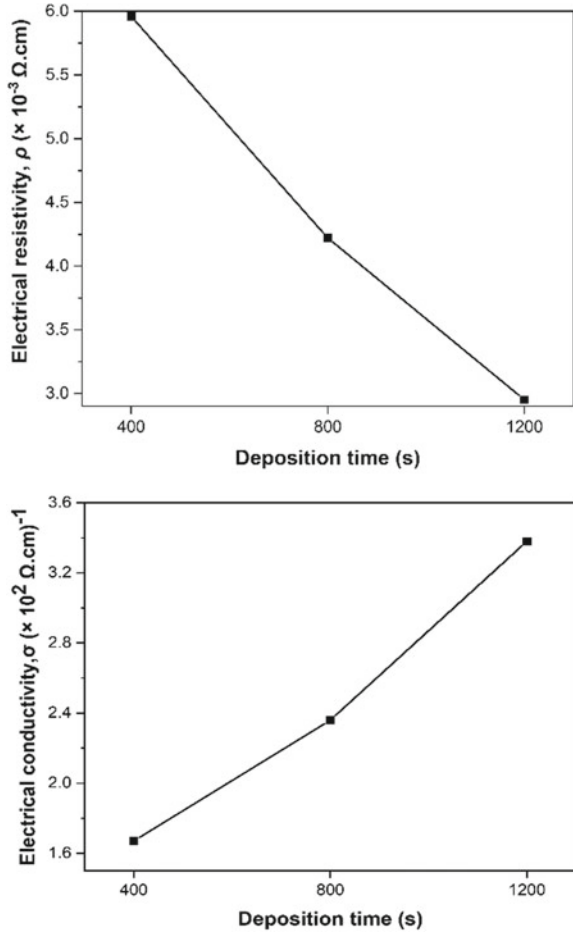


Table 6.1 Electrical resistivity and conductivity of ITO thin films deposited on Si at different deposition time

ITO deposition time (s)	Electrical resistivity, ρ ($\times 10^{-3} \Omega \text{ cm}$)	Electrical conductivity σ ($\times 10^2 \Omega \text{ cm})^{-1}$
400	5.96	1.67
800	4.22	2.36
1200	2.95	3.38

of transmittance. ITO thin film absorbance shows the absorption edges show a good agreement with transmittance data.

Figure 6.3 shows AFM images of ITO thin films deposited on Si. The images show granules present in all samples. As a result, increasing in deposition time of ITO, decreased the surface roughness (R_q). The R_q of ITO thin films decreased from 12.6

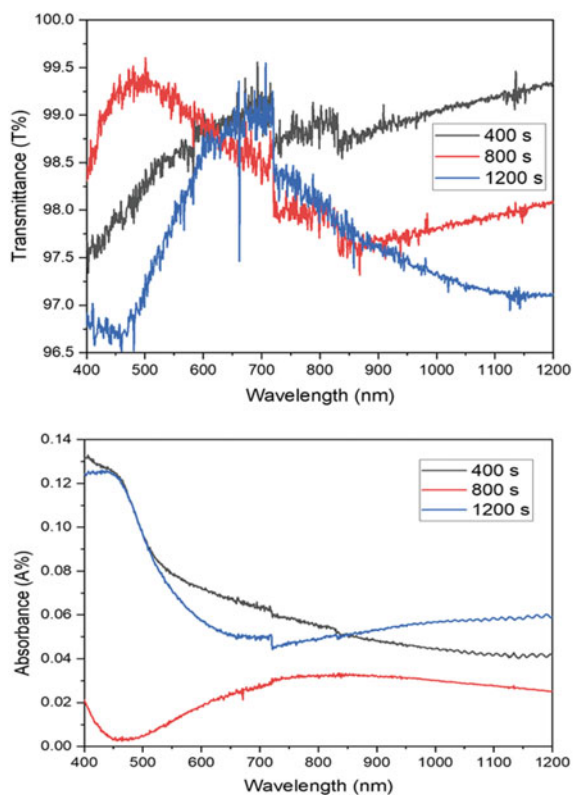


Fig. 6.2 Transmittance and absorbance of ITO thin films deposited on glass substrates against wavelength

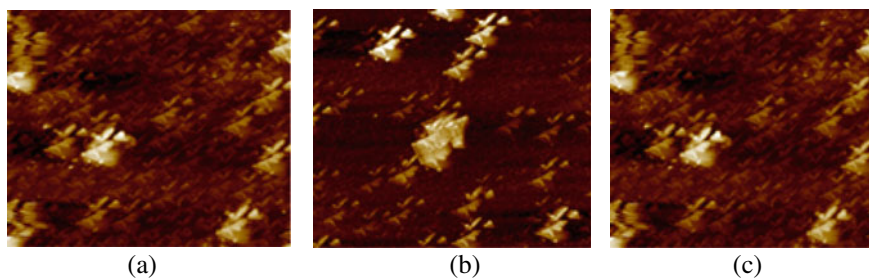


Fig. 6.3 AFM images of ITO thin films deposited on Si substrates at different deposition time **a** 400 s, **b** 800 and **c** 1200 s

to 5.03 nm with the increment of deposition time from 400 to 1200 s, respectively. The 400 s deposition time formed higher R_q compared to other ITO thin films which is 12.6 nm. For ITO thin films at 1200 s deposition time has lower R_q which is 5.03 nm. According to Kumar et al. [16], the R_q increased with thickness, and this is because of the aggregation of grains or increased in crystallinity. However, the results show a decrease of R_q with increasing deposition time. The results achieved may be due to the formation of the coarse grain and dense wrinkles.

6.4 Conclusion

DC magnetron sputtering method used to deposit ITO on Si and glass substrates at different ITO deposition time. This study investigated the influence of different deposition times of ITO thin film on the electrical, optical, and topological properties. All properties strongly depended on the deposition time of ITO. ITO films exhibit low resistivity of $2.95 \times 10^{-3} \Omega \text{ cm}$ and approximately 96% average transmittance in all wavelength range. The high transmittance of ITO thin films, particularly in the visible light region, was a critical parameter for solar cell applications. AFM images showed that the R_q of ITO thin films decreased as the deposition time increased. These results may be due to the formation of the coarse grain, and wrinkles on the surface. One of the important conclusions of study is that ITO thin film can be used as a front contact of solar cells.

Acknowledgements The author(s) would like to thank Universiti Tun Hussein Onn Malaysia for the financial and facilities supported by UTHM TIER grant H845.

References

1. M. Mazur et al., Effect of thickness on optoelectronic properties of ITO thin films. *Circuit world*, vol. ahead-of-print, no. ahead-of-print (2020). <https://doi.org/10.1108/CW-11-2019-0170>
2. P. Prepelita, M. Filipescu, I. Stavarache, F. Garoi, D. Craciun, Transparent thin films of indium tin oxide: morphology–optical investigations, interdependence analyzes. *Appl. Surf. Sci.* **424**, 368–373 (2017)
3. M.S. Siddiqui, A.K. Saxena, S.P. Singh, Deposition and characterization of ITO thin film over glass for defogger application and for solar photovoltaics. *Int. J. Curr. Eng. Technol.* **8**(04) (2018)
4. R. Maniyara et al., Highly transparent and conductive ITO substrates for near-infrared applications. *APL Mater.* **9**(2), 021121 5 (2021)
5. W. Liu, S. Cheng, Photoelectric properties of ITO thin films deposited by DC magnetron sputtering. *J. Semicond.* **32**(1), 013002 (2011)
6. A.K. Isiyaku et al., Preparation of Sn doped In₂O₃ multilayer films on n-type Si with optoelectronics properties improved by using thin Al–Cu metals interlayer films. *Mater. Sci. Semicond. Process.* **131**(105870), 105870 (2021)
7. P. Prepelita et al., Rapid thermal annealing for high-quality ITO thin films deposited by radio-frequency magnetron sputtering. *Beilstein J. Nanotechnol.* **10**, 1511–1522 (2019)

8. A.H. Shukor, H.A. Alhattab, I. Takano, Electrical and optical properties of copper oxide thin films prepared by DC magnetron sputtering. *J. Vac. Sci. Technol. B Nanotechnol. Microelectron.* **38**(1), 012803 (2020)
9. D. Zhang, in *Thermal Barrier Coatings Prepared by Electron Beam Physical Vapor Deposition (EB-PVD)*. Thermal Barrier Coatings (Elsevier, 2011), pp. 3–24
10. O. Stadel et al., Continuous YBCO deposition onto moved tapes in liquid single source MOCVD systems. *Physica C Supercond.* **341–348**, 2477–2478 (2000)
11. H. Soonmin, Thermal evaporation of thin films. *Review* **23**(11), 2695–2699 (2015)
12. A.K. Isiyaku, A.H. Ali, R.A. Ahmad, B.Z. Bhari, Optoelectronic simulation properties of transparent conducting indium tin oxide for solar cell application. *J. Sci. Technol.* **9**(3) (2017)
13. M. Huang, Z. Hameiri, S. Venkataraj, A. Aberle, T. Mueller, Characterization and optimisation of indium tin oxide films deposited by pulsed DC magnetron sputtering for heterojunction silicon wafer solar cell applications. *Energy Procedia* **33**, 91–98 (2013)
14. A.K. Isiyaku, A.H. Ali, N.F. Ramly, Characterization of Cr/Ag metal thin films interaction with infra-red laser. *Mater. Today* **7**, 692–696 (2019)
15. R. Kainikkara Vatakketath, *Investigation of the Transparent Conducting Oxide (TCO) Material Used in CIGS Thin Film Solar Cell in Midsummer AB* (Uppsala Universitet, Sweden, 2020)
16. K.J. Kumar, N.R.C. Raju, A. Subrahmanyam, Thickness dependent physical and photocatalytic properties of ITO thin films prepared by reactive DC magnetron sputtering. *Appl. Surf. Sci.* **257**(7), 3075–3080 (2011)

Chapter 7

The Effect of pH on Particle Size of Hydrothermally Reduced Graphene Oxide



Yi Lin Chan, Fahmiruddin Esa, Kok Yeow You, and Man Seng Sim

Abstract The reduction of graphene oxide (GO) was previously carried out using chemical reducing agents hydrazine, dimethyl sulfoxide (DMSO) and sodium borohydride (NaBH_4). These chemical agents endanger the environment and cause reduced graphene oxide (RGO) sample to form aggregation. In this paper, RGO samples were prepared using hydrothermal synthesis under different pH mediums ($\text{pH} = 4, 7$ and 11). The effects of pH value on surface morphology, chemical compositions and particle sizes were investigated using field emission scanning electron microscope (FESEM), fourier transform infrared (FTIR) and Malvern Zetasizer Nano ZS. FESEM shows RGO samples present in nanosheets and wrinkled structures. FTIR indicates that the characteristic peaks of RGO are in minimum value at different pH values. Particle sizes of RGO samples prepared in different pHs vary from 344.0 nm to 751.1 nm. The particle sizes of RGO samples decrease with increasing of pH value due to electrostatic repulsion on the graphene sheets.

7.1 Introduction

Graphene oxide (GO) is a graphene derivative material that possesses hydroxyl group, epoxide and carbonyl group functionalize on the graphene sheet layer and carbonyl group functionalize with edges of the graphene sheet [1]. GO has been utilized widely in biosensor application owing to its high chemical and electrochemical activity. However, the conjugated network and phi electron conductivity are affected if any oxidation occurs in these groups [2]. Reduced graphene oxide (RGO) is one of the

Y. L. Chan · F. Esa (✉)

Department of Physics and Chemistry, Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, 84600 Pagoh, Muar, Johor, Malaysia

e-mail: fahmir@uthm.edu.my

K. Y. You · M. S. Sim

School of Electrical Engineering, Faculty of Engineering, Universiti Teknologi Malaysia, 81310 Skudai, Johor Bahru, Johor, Malaysia

e-mail: kyyou@fke.utm.my

best candidates among the graphene derivatives because of its large surface areas made up of a huge number of residual electroactive sites and structures similar to graphene [3].

Reduction of GO can be prepared using hazardous reducing agents hydrazine, dimethylhydrazine, dimethyl sulfoxide (DMSO), sodium borohydride (NaBH_4), high temperature reaction or a combination of both reducing agent and high temperature reaction [4, 5]. Besides that, the use of hydrazine, dimethylhydrazine, DMSO and NaBH_4 causes RGO samples to aggregate in reduction reaction [6]. A reduction reaction held at high temperature ($900\text{ }^\circ\text{C}$) produces several by-products including carbon monoxide, carbon dioxide and water by depleting the carbon backbone of GO [7]. Additionally, thermal reduction reaction usually requires heating in the presence of argon or hydrogen gases that makes the reaction to be a difficult process [8].

To our knowledge, several studies reported on reduction of GO using chemical reducing agent, high temperature or a combination of both chemical and high temperature. The use of hazardous hydrazine, DMSO and NaBH_4 are not environmental friendly method and bring about the aggregation of RGO samples. Besides that, chemical reduction method using NaBH_4 involves tedious long hour process of sonication (2 h) and stirring (12 h). Hence, it is worthwhile to explore an environmental friendly and simple preparation method. In this work, hydrothermal method was deployed for reduction of GO by varying pH values. Sodium hydroxide and hydrochloric acid were used to change the pH of the medium to become pH 4, pH 7 and pH 11. The effects of pH value on the surface morphology, chemical composition and particle size of RGOs were studied using FESEM, FTIR and Malvern Zetasizer Nano ZS.

7.2 Experimental Setup

7.2.1 Material

Graphene oxide powder was purchased from SkySpring Nanomaterials. Ethanol ($\text{C}_2\text{H}_6\text{O}$) and Sodium Hydroxide (NaOH) were bought from Merck. Hydrochloric acid (HCL) was obtained from R&M Malaysia.

7.2.2 Preparation of RGO

All RGO samples prepared under the same concentration with different pH values. RGO was synthesized using 1.0 g GO powder dispersed in 150 ml deionized water. The GO suspension was sonicated for 30 min and stirred for 30 min. Then, 1.0 M NaOH or 2% HCL was added dropwise to the suspension until its pH became 4, 7 and 11 under constant stirring. The suspension was transferred to a Teflon lined

autoclave for hydrothermal treatment at 180 °C for 8 h. The obtained product was centrifuged with ethanol for several times. Finally, it was dried in an oven at 100 °C for 24 h. The hydrothermally prepared RGO was labelled as RGO-A, RGO-N and RGO-B with the different mixture conditions in pH 4 (acidic), pH 7 (neutral) and pH 11 (basic) respectively.

7.2.3 Material Characterisation

RGO powders obtained from hydrothermal synthesis were palleted using hydraulic pressure machine under the pressure 1.4 tons for 3 min. Field emission scanning electron microscope (FESEM) with EDS (JEOL JSM-7600F) was used to investigate surface morphology of RGO pellets. Fourier transform infrared (FTIR) analysis was obtained via FTIR spectroscopy (Perkin Elmer Spectrum 100) in transmission mode over wavenumber range 450–4000 cm^{-1} . Particle sizes of RGO suspensions were determined using Malvern Zetasizer Nano ZS.

7.3 Results and Discussions

7.3.1 Morphological Analysis Using FESEM

Figure 7.1 shows the morphology of RGO prepared in different pH mediums. It can be observed that the all the RGOs are made up of sheet layers. These curved and wrinkled structures prevent the graphene sheets from forming stacked layers among graphene sheet layers [9] and agglomerate of the RGO samples arising from Van der Waals force [10]. Additionally, these curved and wrinkled morphologies endow mesoporous nature of RGO [9]. These mesopores may trap carbon dioxide gas released during hydrothermal reaction of RGO.

7.3.2 Functional Group Analysis Using FTIR

The FTIR spectra of RGO obtained for pH 4, pH 7 and pH 11 are presented in Fig. 7.2. The IR peaks of RGO samples are C–H bending vibration (667 cm^{-1} and 868 cm^{-1}) [10], C–O stretching vibration (1047 cm^{-1} and 1217 cm^{-1}) [11], O–H stretching vibration (1365 cm^{-1} , 2661 cm^{-1} and 2258 cm^{-1}) [12], C=O stretching vibration (1740 cm^{-1}) [13]. Additional peaks can be seen at around 2300 cm^{-1} and can be related to the carbon dioxide molecule originated from the RGO [14]. All these characteristic peaks corresponding to RGO samples are quite similar. Similar

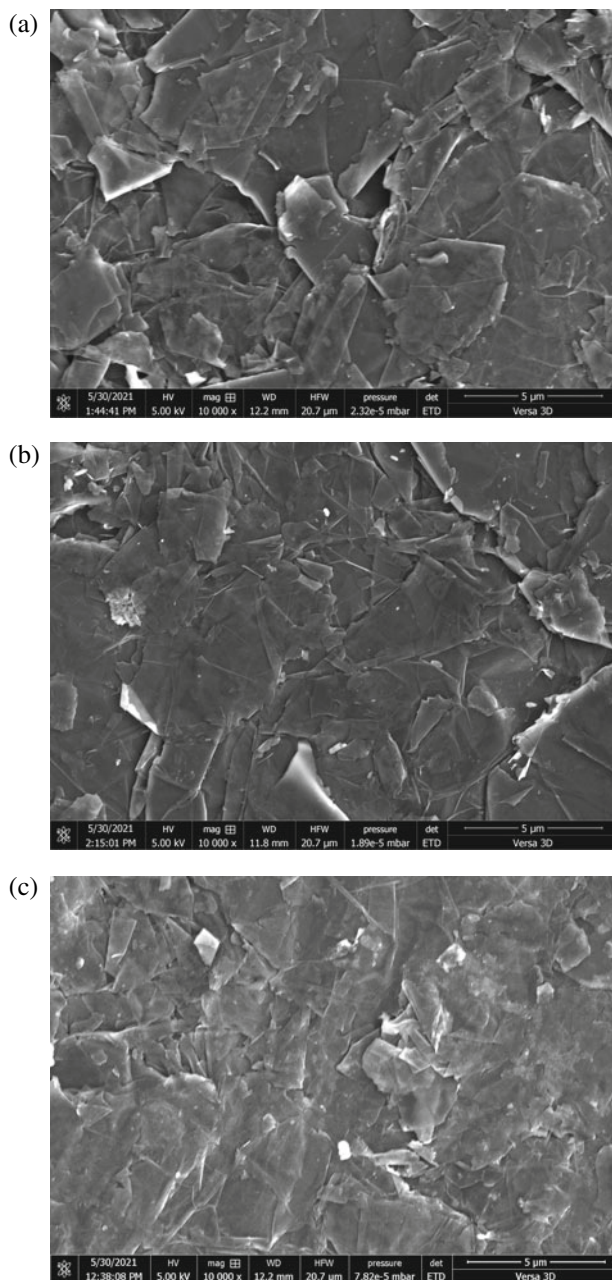


Fig. 7.1 FESEM images of **a** RGO-A, **b** RGO-N and **c** RGO-B hydrothermally prepared at pH 4, pH 7 and pH 11 respectively

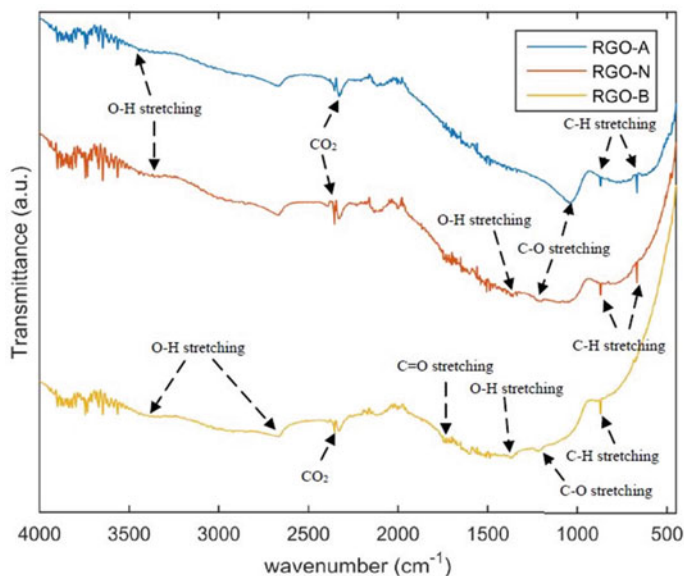


Fig. 7.2 FTIR spectra of hydrothermally prepared RGO samples at different pH mediums (4, 7 and 11)

finding were obtained from the previous studies reporting on the FTIR spectra of preparation of RGO samples with increasing the pH in the medium [6, 10, 15, 16].

7.3.3 Malvern Zetasizer Nano ZS

Dynamics light scattering (DLS) is a suitable technique to estimate the size of spherical-shaped particles based on the particle movement in liquid. However, RGO samples possess a large ratio of length ranging from few microns to thickness ranging from few nanos [1]. Hence, the DLS results of as-mentioned samples are presumed to be close to the side dimension of RGO samples. Figure 7.3 shows the particle size distribution of RGO measured via DLS. The particle sizes among RGO samples distributed normally without positively and negatively skewed. The particle size distribution and its average value among RGO with different pHs are obtained from Zeta software and tabulated in Table 7.1.

The effect of various pHs (4, 7 and 11) on particle size of RGO samples is investigated and displayed in Fig. 7.4. The average particle size as shown in Table 7.1 was utilized because it shows the most population of particle size in respect of each RGO samples. The population decreased to 548.0–344.0 nm after the RGO was added NaOH dropwise until the pH of the medium become 11 in the sample preparation. Noticeably, the particle size of RGO-B decreased in the basic medium

Fig. 7.3 Particle size distribution of RGO samples at **a** pH 4, **b** pH 7 and **c** pH 11

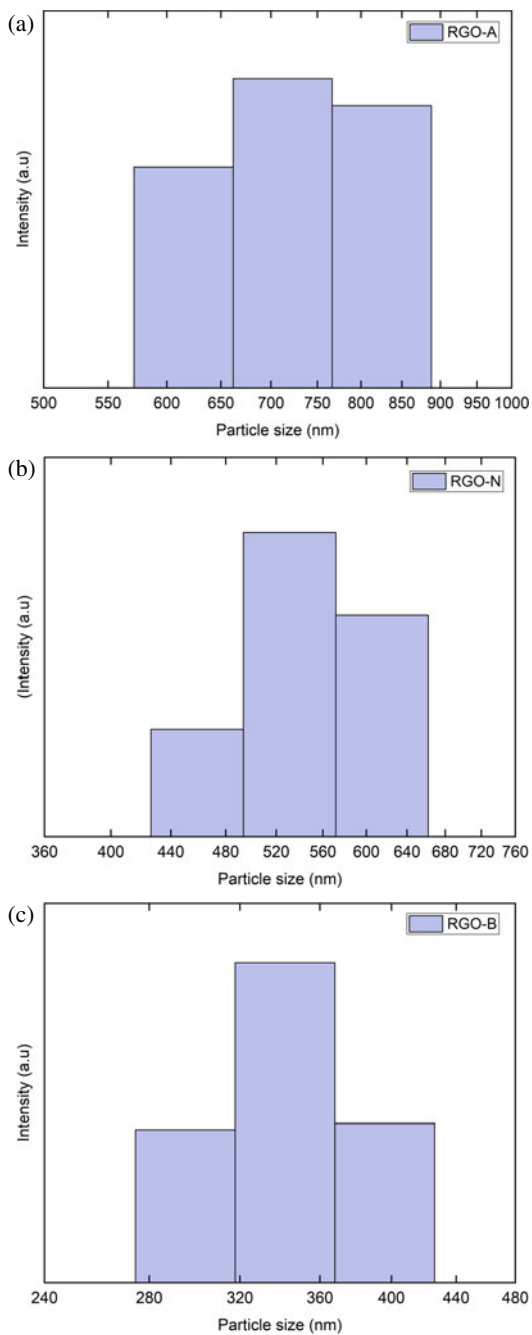
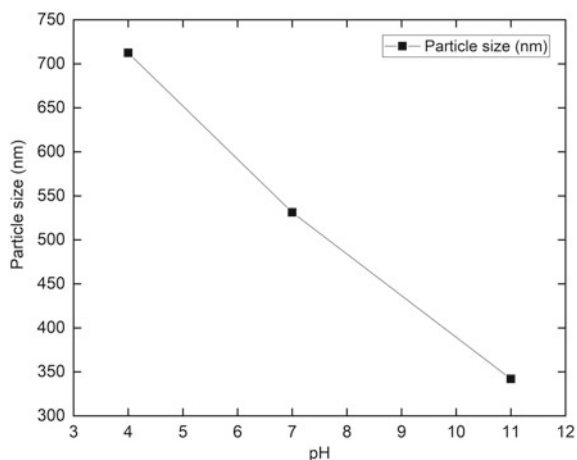


Table 7.1 The particle size of RGO samples in different pHs

Samples	pH values	Average particles size (nm)	Range of particle size (nm)
RGO-A	4	751.1	615.1–825.0
RGO-N	7	548.0	458.7–615.1
RGO-B	11	344.0	295.3–396.1

Fig. 7.4 The dependency of particle sizes of RGO samples on various pHs (4, 7 and 11)

as compared to that of neutral medium. This can be related to the increasing pH value lowers the tendency of graphene sheets to agglomerate them. NaOH produces negative charge hydroxide ions (OH^-) in solution that generates electrostatic repulsion among the graphene sheets and thus prohibits the agglomeration of the graphene sheets [17]. Besides that, another plausible reason for this decreasing particle size in basic medium is NaOH serves as a hydrogenating agent for GO and stabilize the GO colloid [18]. The formation of activated graphene sheets begins by the ability of NaOH to remove oxygen atom from the surface of GO [19, 20]. These activated graphene sheets start to stabilize among them and reduce their effective sizes in the basic solution. Therefore, this phenomenon contributes to smaller sizes of activated graphene sheets. On the contrary, a larger size of RGO-A can be observed in 751.1 nm. This can be ascribed to the increase of concentration of hydrogen ions (H^+) in the solution which gives rise to the increase in size of graphene sheet. HCL dissociates into H^+ when it is dissolved in solution. The formation of H^+ ions decreases the pH value of the solution and increases protonation of both C–OH and COOH acidic groups existing on the graphene sheets. The protonation process reduces the negative charges that acts as electrostatic repulsion on the graphene sheets and thus finally bring about the agglomeration of graphene sheets [21]. Kashyap et al reported on the effect of pH value on chemically modified graphene oxide over a range of 2–12 with the increment of pH value by 1. The reported particle sizes decreased with increasing pH values that agree well with this study [21].

7.4 Conclusion

In this work, RGO samples were prepared using hydrothermal synthesis in different pH mediums pH 4, pH 7 and pH 11. Wrinkle and curved structures can be observed in FESEM images for all RGO samples. FTIR spectra depict the characteristic peaks of RGO samples present in minimum intensity, showing the reduction of GO. The role of pH greatly affects the particle size of RGO samples. The particle size increases in acidic medium (751.1 nm) as compared to that of basic medium (344.0 nm) due to the concentration of H⁺ ions increases and favours protonation process. This process decreases the electrostatic repulsion of the graphene sheets and thus results in agglomeration of graphene sheets.

References

1. A.T. Smith, A.M. LaChance, S. Zeng, B. Liu, Sun, L., *Nano Mater. Sci.* **1**(1), 31 (2019)
2. J. Zhao, S. Pei, W. Ren, L. Gao, H.M. Cheng, *ACS Nano* **4**(9), 5245 (2010)
3. D.F. Báez, H. Pardo, I. Laborda, J.F. Marco, C. Yáñez, S. Bollo, *Nanomater* **7**(7), 1 (2017)
4. S. Stankovich, D.A. Dikin, R.D. Piner, K.A. Kohlhaas, A. Kleinhammes, Y. Jia, Y. Wu, S.T. Nguyen, R.S. Ruoff, *Carbon* **45**(7), 1558 (2007)
5. Y. Wu, S.T. Nguyen, R.S. Ruoff, *Carbon* **45**(7), 1558 (2007)
6. G. Tai, T. Zeng, H. Li, J. Liu, J. Kong, F. Lv, *Mater Res Express* **1**(3), 035605 (2014)
7. D. Luo, G. Zhang, J. Liu, X. Sun, *J. Phys. Chem. C* **115**(23), 11327 (2011)
8. P.P.A. Jose, M.S. Kala, N. Kalarikkal, S. Thomas, *Mater. Today Proc.* **5**(8), 16306 (2018)
9. Y. Bai, R.B. Rakhi, W. Chen, H.N. Alshareef, *J. Power Sources* **233**, 313 (2013)
10. E.C. Vermisoglou, T. Giannakopoulou, G. Romanos, M. Giannouri, N. Boukos, C. Lei, C. Lekakou, C. Trapalis, *Appl. Surf. Sci.* **358**, 100 (2015)
11. Z. Lin, Ph. D. dissertation, Georgia Institute of Technology (2014)
12. F.T. Johra, W.G. Jung, *Appl. Surf. Sci.* **357**, 1911 (2015)
13. L. Shahriary, A.A. Athawale, *Int. J. Renew. Energy Environ. Eng.* **2**(01), 58 (2014)
14. I.O. Faniyi, O. Fasakin, B. Olofinjana, A.S. Adekunle, T.V. Oluwasusi, M.A. Eleruja, E.O.B. Ajayi, *Appl. Sci.* **1**(10), 1 (2019)
15. C. Bosch-Navarro, E. Coronado, C. Martí-Gastaldo, J.F. Sánchez-Royo, M.G. Gómez, *Nanoscale* **4**(13), 3977 (2012)
16. D.E. Glass, G.K. Surya Prakash, *Electroanalysis* **30**(9), 1938 (2018)
17. Y. Zhou, Q. Bao, L.A.L. Tang, Y. Zhong, K.P. Loh, *Chem. Mater.* **21**(13), 2950 (2009)
18. T. Kuila, P. Khanra, N.H. Kim, J.K. Lim, J.H. Lee, *J. Mater. Chem. A* **1**(32), 9294 (2013)
19. Y. Zhu, S. Murali, M.D. Stoller, K.J. Ganesh, W. Cai, P.J. Ferreira, A. Pirkle, R.M. Wallace, K.A. Cychoz, M. Thommes, D. Su, E.A. Stach, R.S. Ruoff, *Science*, **332**(6037), 1537 (2011)
20. D. Zhang, X. Zhang, Y. Chen, C. Wang, Y. Ma, *Electrochim. Acta* **69**, 364 (2012)
21. S. Kashyap, S. Mishra, S.K. Behera, *J. Nanopart.* **1** (2014)

Chapter 8

A Comparative Study of Cr and Mo Ultrathin Films as Transparent Conductors in Solar Cell Application



Abdelbaki Hacini, Ahmad Hadi Ali, and Nurul Nadia Adnan

Abstract This research investigates the properties of chromium Cr and molybdenum Mo ultra thin films deposited on silicon and glass substrate using an RF sputtering system. This study helps to make a new efficient, transparent conductor for the solar cell application. These thin films exhibit a high optical transmittance belongs to the layer with 5 nm of thickness which attained $> 85\%$ of transmittance in the spectral range 400–800 nm. The surface morphology of Cr and Mo thin film varies with thickness, where the grain size and surface roughness increase with thickness. Additionally, the electrical properties were measured using the four-point probe. The lower resistivity obtained was $5.56 \times 10^{-4} \Omega \cdot \text{cm}$ and $6.87 \times 10^{-4} \Omega \cdot \text{cm}$ at 30 nm of thickness for Cr and Mo thin film, respectively. The main contribution of this work was to obtain a thin film with low resistivity and high transparency for the transparent conducting application. These enhanced properties will reduce the optical and electrical losses caused by the light reflection and the electric contact.

8.1 Introduction

The realisation of suitable transparent conducting materials with low cost is a significant challenge for researchers. These materials are essential for all electronic and optoelectronic devices [1–4]. Up to now, the transparent conducting oxides thin films are dominated for the transparent conducting applications. These oxides combine the high electrical conductivity with the high transmittance in the visible and near-infrared ranges [4–8]. However, they have some disadvantages, including the need

A. Hacini (✉) · A. H. Ali · N. N. Adnan

Laser and Semiconductor Technology Research Group, COR PDSR, Department of Physics and Chemistry, Faculty of Applied Sciences and Technology, Pagoh Educational Hub, Universiti Tun Hussein Onn Malaysia, 84600 Pagoh, Johor, Malaysia

A. H. Ali

e-mail: ahadi@uthm.edu.my

N. N. Adnan

e-mail: nadia@uthm.edu.my

for postdeposition or annealing treatment after deposition, the toxic nature of some oxides, and the high production costs of these materials [9]. These disadvantages lead to the search for an alternative material. In this context, the recent development in thin metal films exhibits the ability to replace transparent oxides at very low thicknesses because they are optically transparent and have very low resistivity [9]. Ultrathin films are thin films that have a shallow thickness (less than 100 nm). This category of thin films are used to improve the absorbance of the solar cell and decrease the reflection that occurs on the top surface. Moreover, it helps to boost the mobility of electrons due to the high carriers concentration. In this matter, chromium and molybdenum were chosen because of their good conductivity and high transparency at the low scale of thickness.

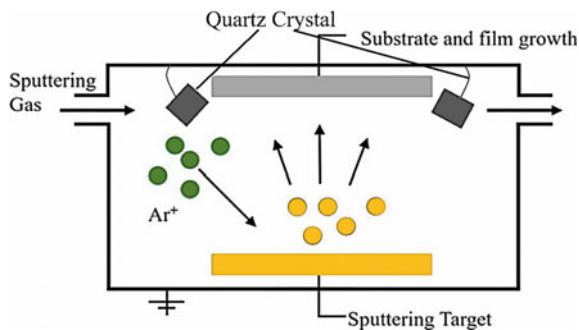
In this respect, it is common knowledge that thin films' properties are influenced by other properties such as the deposition parameters [10–13], the thickness [14], the existence of oxygen [5], and doping carrier [15]. Consequently, the investigation into thin metal films' thickness plays a significant role in understanding the growing processes of thin films and determining a new shape for thin films. These two aspects will enhance the optical and morphological properties to be applicable to optoelectronic devices. Nowadays, the sputtering deposition method has a great place in fabricating thin metal films because of its advantages, including a low-temperature deposition, good reproducibility, and the most important is the high growth rate [9, 16].

This work presents a comparison of chromium and molybdenum ultrathin films optically, electrically, and morphologically concerning the change in film thickness for transparent conducting application.

8.2 Methodology

The chromium Cr and molybdenum Mo thin film were deposited onto glass and n-type silicon Si substrates using an RF magnetron sputtering machine. The glass substrates are used to measure the optical properties, where the Si substrates are used to identify thin films' morphological and electrical properties. Both the glass and silicon substrate were sputtered in a vacuum chamber under high pressure 10^{-4} Pa. The glass substrates were cut and cleaned using the Decon-90 glass cleaner. Meanwhile, the Si pieces were cleaned using an ultrasonic cleaner where the samples were immersed in acetone solution, propanol, and DI water for 10 min. Then, these glass and Si pieces were dried using nitrogen gas and placed under RF waves for 10 minutes in a plasma cleaner. The Cr and Mo films were deposited for different thicknesses from 5 to 30 nm with 5nm. However, the thickness was controlled by the two quartz crystals integrated within the chamber. During the deposition, the mass of material grows on the surface of quartz crystal and causes fluctuation in the frequency. At the same time, the quartz crystal measures the change in the mass per crystal area by measuring the change of frequency (Fig. 8.1).

Fig. 8.1 A schematic diagram of The experimental setup



The characterisation of samples included optical, electrical, and morphological. The electrical properties were determined using the four-point probe system. While the optical properties were measured using a UV-Vis spectrometer (UV-3600i Plus, SHIMADZU) in the range from 300 to 800 nm. Moreover, the AFM (Bruker Dimension Edge atomic force microscope) was used to obtain the surface topology.

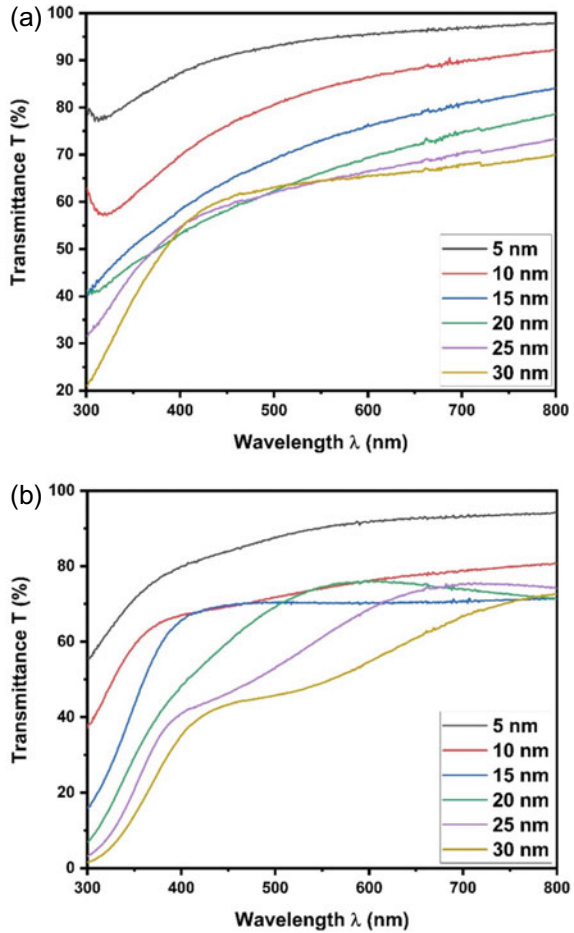
8.3 Results and Discussion

Figure 8.2 shows the optical transmittance spectra of Cr and Mo thin films for different thicknesses varying from 5 nm to 30 nm in the wavelength range of 300–800 nm. The spectrum shows an average of transmittance above 80% transmittance for 5 nm thickness of Mo and Cr thin films at a wavelength of at least 400 nm. However, the transmittance decreases for the same thickness (=5nm) in the UV range (300–400) above 70%. Both materials exhibit a slight decrease in transmittance with thin-film thickness because of surface scattering caused by the carriers. The increase of the thickness will boost the electrons on the surface that influence inter-band transitions and occur more absorption of light. In comparison, Cr thin films show a high transmittance than Mo thin films.

Figures 8.3 and 8.4 the three-dimensional (3D) AFM images of Cr and Mo thin films measured in an area of $5.0 \times 5.0 \mu\text{m}^2$ for different thicknesses. These figures show different morphologies for each material with varying thicknesses. Practically, the deposition technique with different parameters has a significant impact on surface morphology. It can be observed in the 3D images that the surface morphology contains a series of separate peaks. Moreover, the size and distribution of grains change with the thin film thicknesses. Figure 8.5a exhibits the root mean square (RMS) roughness varying with thin-film thickness in the range of 5–30 nm. The RMS roughness increases from 0.95 nm to 1.45 nm for the Cr thin film meanwhile, the RMS roughness of the Mo thin film increase from 0.71 nm to 1.09 nm.

The general trending of the RMS roughness is increasing over the 5–30 nm thickness despite having decreased values for Mo thin film in 15 nm and 20 nm thickness. Similarly, the RMS roughness of Cr thin film decreases in thickness (10 nm, 20 nm,

Fig. 8.2 The transmittance spectra of glass coated with **a** Cr and **b** Mo thin films



and 25 nm). The roughness is associated with the agglomeration and the distribution of the particles on the surface. The increase of thin-film thickness will increase the number and the size of particles that cause an increase in the roughness [16, 17]. However, the decrease in roughness of Mo thin film at 20 nm and Cr thin film at 10 nm and 25 nm due to the deposition process improves the surface quality and minimises the roughness over a longer time.

Figure 8.5b shows the growth of the grain size of Cr and Mo thin film as a function of thickness. The grain size increases from 56.2 nm to 151.5 for Cr thin film and from 55.9 nm to 110.5 nm for Mo thin film. This trend is due to the thin-film thickness and the deposition process. The sputtering process will condense atoms into the various separated nucleus at small thicknesses, then the agglomeration of the nucleus will form the 3D islands or particles [18]. In comparison, Cr thin films show a bigger

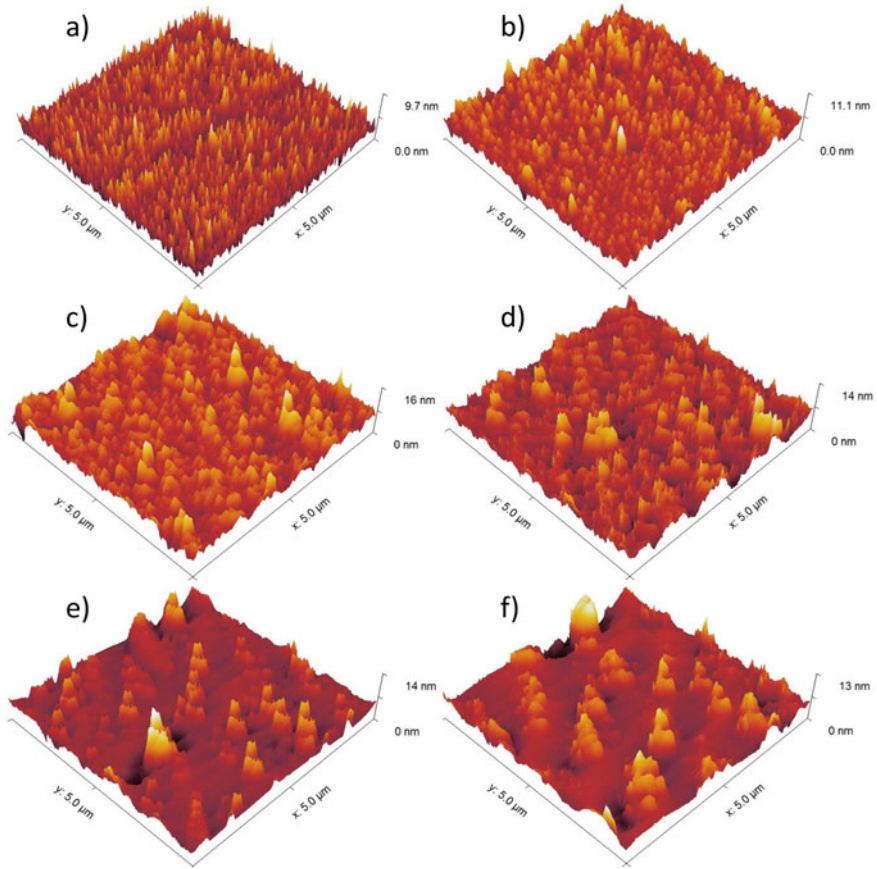


Fig. 8.3 AFM images of thin Cr films for the as-deposited on silicon **a** 5 nm **b** 10 nm **c** 15 nm **d** 20 nm **e** 25 nm **f** 30 nm

grain size than that of Mo thin films. While Mo thin exhibit a lower roughness than the Cr thin film.

Table 8.1 displays the variation of resistivity of Cr and Mo films with thickness. The results indicate that the resistivity of both materials decreases with increasing the film thicknesses. The decrease in resistivity of the Cr and Mo films is due to the increase in grain size, which boosts the mobility of carriers inside the thin film [19]. The growth of grain size decreases in grain boundary and defects on the thin film surface, which reduces potential barriers for electron mobility. Cr exhibits a lower resistivity as compared with Mo; correlating with the grain size of Cr is more significant than that of Mo [19].

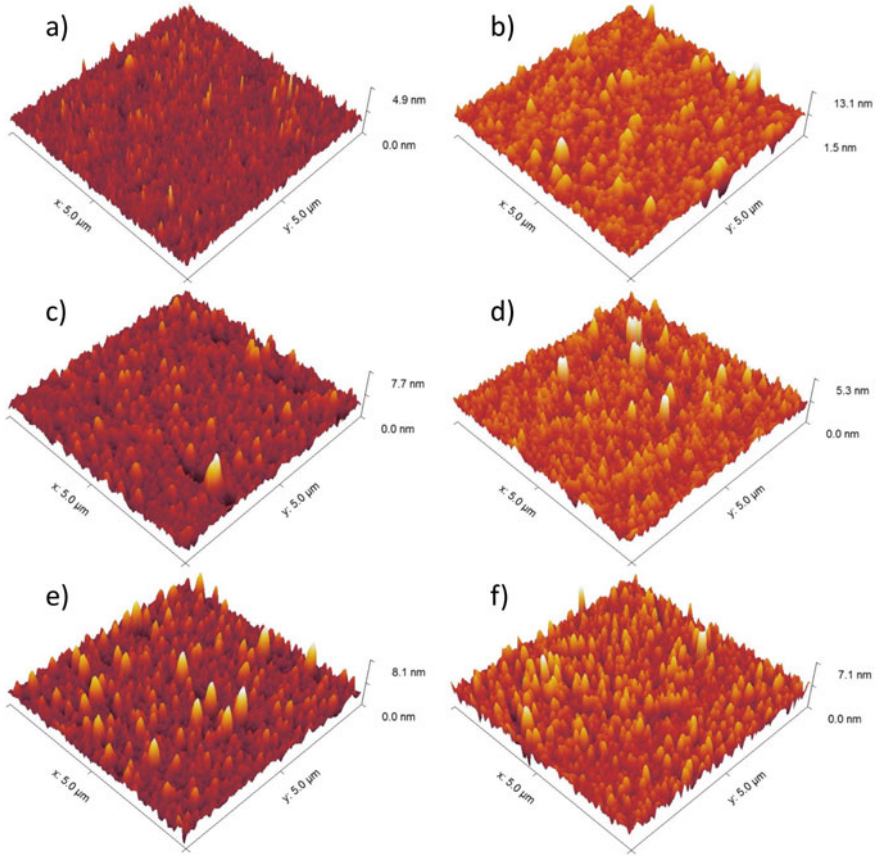


Fig. 8.4 AFM images of thin Mo films for the as-deposited on silicon **a** 5 nm **b** 10 nm **c** 15 nm **d** 20 nm **e** 25 nm **f** 30 nm

8.4 Conclusion

In this work, the Cr and Mo thin films were deposited on glass and Si substrates using RF magnetron sputtering. The influence of thickness on the optical, morphological, electrical properties, and elemental composition of Cr and Mo thin films were investigated. The optical transmittance spectra show high transmittance in the visible and near-infrared ranges above 80% for 5 nm of thickness. AFM results show an increase of the surface roughness similarly to grain size due to the sputtering process. The electrical resistivity decrease with the increase of thickness. The lower resistivity obtained was $5.56 \times 10^{-4} \Omega \cdot \text{cm}$ and $6.87 \times 10^{-4} \Omega \cdot \text{cm}$ for Cr and Mo thin film, respectively.

Fig. 8.5 Cr, MO grain size, and roughness measured of thin films deposited

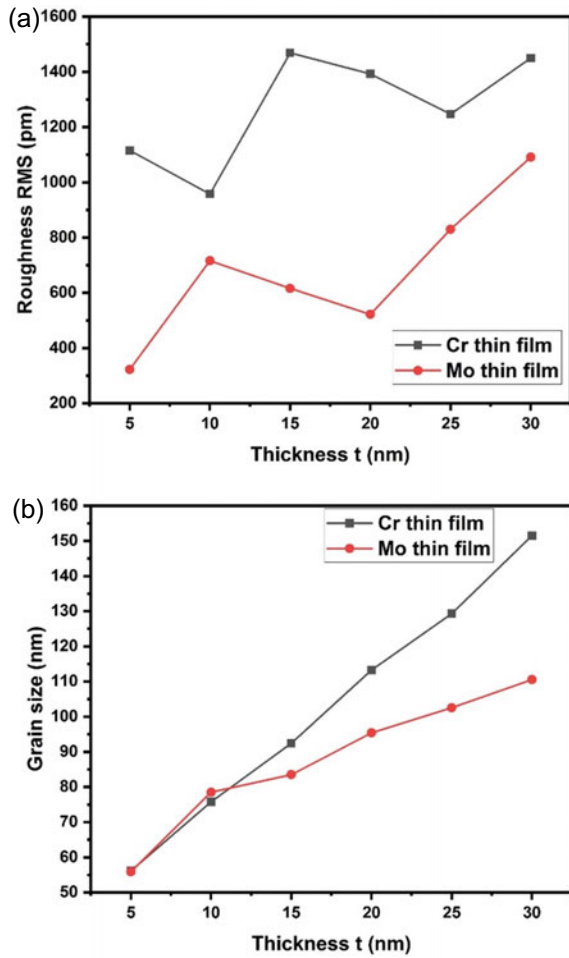


Table 8.1 Cr, Mo resistivity measured of thin films deposited on silicon

Thickness (nm)	10	15	20	25	30
Resistivity of Cr film $\times 10^{-4}$ ($\Omega \cdot \text{cm}$)	11.6	9.58	8.29	7.45	5.56
Resistivity of Mo film $\times 10^{-4}$ ($\Omega \cdot \text{cm}$)	12.12	10.5	9.73	8.37	6.87

Acknowledgements The author(s) would like to thank the Government of Malaysia and Universiti Tun Hussein Onn Malaysia for funding this research project through the UTHM TIER1 Grant (H845) and FRGS RACER Grant (K166).

References

1. H.E. Shin, C.O. Jeong, J. Song, Metal Lines and ITO PVD. *Flat Panel Disp. Manuf.* 193–208 (2018)
2. S.-H. Jeong et al., Universal high work function flexible anode for simplified ITO-free organic and perovskite light-emitting diodes with ultra-high efficiency. *NPG Asia Mater.* **9**(7), e411–e411 (2017)
3. A.K. Isiyaku, A.H. Ali, N. Nayan, Structural optical and electrical properties of a transparent conductive ITO/Al-Ag/ITO multilayer contact. *Beilstein J. Nanotechnol.* **11**(September), 695–702 (2020). <https://doi.org/10.3762/bjnano.11.57>
4. A.H. Ali, Z. Hassan, A. Shuhaimi, Enhancement of optical transmittance and electrical resistivity of post-annealed ITO thin films RF sputtered on Si. *Appl. Surf. Sci.* **443**, 544–547 (2018). <https://doi.org/10.1016/j.apsusc.2018.03.024>
5. J. Gwamuri, M. Marikkannan, J. Mayandi, P.K. Bowen, J.M. Pearce, Influence of oxygen concentration on the performance of ultra-thin RF magnetron sputter deposited indium tin oxide films as a top electrode for photovoltaic devices. *Materials (Basel)* **9**, 63 (2016)
6. H. Taha et al., Probing the effects of thermal treatment on the electronic structure and mechanical properties of Ti-doped ITO thin films. *J. Alloys Compd.* **721**, 333–346 (2017)
7. M.F. Al-Kuhaili, Electrical conductivity enhancement of indium tin oxide (ITO) thin films reactively sputtered in a hydrogen plasma. *J. Mater. Sci. Mater. Electron.* **31**(4), 2729–2740 (2020). <https://doi.org/10.1007/s10854-019-02813-9>
8. L. Voisin, M. Ohtsuka, S. Petrovska, R. Sergiienko, T. Nakamura, Structural, optical and electrical properties of DC sputtered indium saving indium-tin oxide (ITO) thin films. *Optik (Stuttg)* **156**, 728–737 (2018). <https://doi.org/10.1016/j.ijleo.2017.12.021>
9. K.V. Rajani, S. Daniels, P.J. McNally, F.O. Lucas, M.M. Alam, Ultrathin chromium transparent metal contacts by pulsed dc magnetron sputtering. *Phys. status solidi* **207**(7), 1586–1589 (2010)
10. S. Yang et al., Influence of base pressure on property of sputtering deposited ITO film. *J. Mater. Sci. Mater. Electron.* **30**(14), 13005–13012 (2019)
11. S. Yang et al., Effect of ITO target crystallinity on the properties of sputtering deposited ITO films. *Ceram. Int.* **46**(5), 6342–6350 (2020)
12. O. Boussoum, M.S. Belkaid, C. Renard, G. Halais, F. Farhati, Effect of the annealing gas and RF power sputtering in the electrical, structural and optical properties of ITO thin films (2019)
13. H. Che, M. El Bouanani, Effects of argon sputtering and UV-ozone radiation on the physico-chemical surface properties of ITO. *Nucl. Instrum. Methods Phys. Res. Sect. B Beam Interact. Mater. Atoms*, **414**, 170–175 (2018)
14. J.-H. Kim et al., The effects of film thickness on the electrical, optical, and structural properties of cylindrical, rotating, magnetron-sputtered ITO films. *Appl. Surf. Sci.* **440**, 1211–1218 (2018)
15. Y. Shigesato, D.C. Paine, Study of the effect of Sn doping on the electronic transport properties of thin film indium oxide. *Appl. Phys. Lett.* **62**(11), 1268–1270 (1993)
16. W.-J. Chen, W.-L. Liu, S.-H. Hsieh, Y.-G. Hsu, Synthesis of ZnO: Al transparent conductive thin films using Sol-gel method. *Procedia Eng.* **36**, 54–61 (2012). <https://doi.org/10.1016/j.proeng.2012.03.010>
17. R. Bel Hadj Tahar, T. Ban, Y. Ohya, Y. Takahashi, Tin doped indium oxide thin films: Electrical properties. *J. Appl. Phys.* **83**(5), 2631–2645 (1998)
18. M. Nie, K. Sun, D.D. Meng, Formation of metal nanoparticles by short-distance sputter deposition in a reactive ion etching chamber. *J. Appl. Phys.* **106**(5), 54314 (2009)
19. J. Yang et al., Effect of RF power and substrate temperature on the properties of boron and gallium co-doped ZnO films. *Mater. Sci. Semicond. Process.* **53**, 84–88 (2016)

Chapter 9

Simulation on Optical Absorption for Amorphous Silicon Thin Film Solar Cell with CdSe/ZnS Quantum Dots



Mirza Basyir Rodhuan, Rosmila Abdul-Kahar,
and Amira Saryati Ameruddin

Abstract The light absorption peaks of the cadmium selenide/zinc sulphide (CdSe/ZnS) core/shell spherical quantum dots (QDs) are different, proportional to their sizes. In this study, the various CdSe/ZnS core diameters which were 2.5 nm, 4.0 nm, and 7.0 nm that applied to the amorphous thin film silicon solar cell. The models were designed and simulated using the Finite Element Method simulation software. The aim was to investigate the light absorption by each model design and determine the factors that influence the light absorption. The solar cell models; amorphous silicon solar cell (aSiSC) and amorphous silicon quantum dots solar cell (aSiQDSC) with different core diameters of the QD were determined. The light absorption increased when applying a QD monolayer to the aSiSC model. The light absorption peaks by the aSiQDSC models rose and occurred at the shorter wavelength when the core diameter of the QD decreased. The factors that affect the light absorptions were multi-exciton generations within the QD that generated more photons, the classification of the QDs, and the size of the QDs whereabouts involved tuning the QDs' energy bandgap.

9.1 Introduction

Non-harmful renewable energy is desired to preserve the environment health, for example solar cells [1–3] that undergo the most fundamental process called the light-matter interaction [4]. The quantum technologies and solid state-physics knowledge were expanded and intercepted, resulting the involvement of quantum dot (QD), a particle linked between light and matter [5]. A nanostructure crystalline with semiconductor behaviours analogues to an atomic structure and optical properties; thus,

M. B. Rodhuan · R. Abdul-Kahar (✉) · A. S. Ameruddin
Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, Educational Hub Pagoh, 84600 Muar, Johor, Malaysia
e-mail: rosmila@uthm.edu.my

A. S. Ameruddin
e-mail: amira@uthm.edu.my

the QDs are an artificial atom [6–10]. The multi-body nature of QDs enhanced the strength of the light-matter interaction compared to the atomic emitters; therefore, the QDs can be practical to the solar cells [5]. Previously, some industries that used conventional single-junction solar cells faced a problem known as high-cost cell but low electrical delivery. The absorber of the solar cells was able to absorb lights within a specific region of the solar spectrum. It means that some photons were not being absorbed, and some photons were converted as heat [7, 11].

The presence of the QDs with tuneable bandgap properties depends on their sizes applied to the solar cells, also known as third-generation solar cells, which promise high efficiency, absorption, and low cost [12–14]. The QDs with their tuneable bandgap can harvest more solar spectrum because the QDs have a sizeable intrinsic dipole moment causing rapid charge separation, which can produce three electrons per photon due to multiple exciton generation processes (MEG) [13, 15, 16]. The quantum dots solar cells (QDSC) could theoretically enhance their power efficiency from 20 to 65% [7]. The smaller QDs are giving higher energy confinement and having a shorter wavelength and vice versa. The QDs are changeable in their optical properties as their sizes and shapes are different [2, 7, 17, 18]. However, a photobleaching process is a challenge for the QDs where the optical excitations are irreversible, causing the fluorescence intensity to decrease, so the QDs are blinking. This effect occurred because the photoexcited carriers in the QDs are trapping and de-trapping, which makes the QDs fluctuate between emissive and non-emissive states.

The effect of the photo blinking can be reduced by growing a shell around the QDs then become the core/shell QDs [2, 17]. The crystalline shell around the core is more tuneable, and the core energy level can be manipulated [19]. The shell acted as the separator between the core and the surrounding medium to avoid surface defects [17] and maintain the luminescence properties of the QDs [20].

Nowadays, experimentalists have studied nanostructure areas, but some formulations and theories have deviated as the theoretical approach is not strong enough [21]. Researchers used simulations by using several software and methods to design and handle the experiment computationally to obtain the data for the analysis [22]. Recently, some researchers stated that the investigation for the QDs can be from the experiments and the simulations. This is because the QDs have small dimensions down to nanoscales [20]. The research was stimulated by Cartar [21] using the Finite Element Method (FEM) as a solver to the model of QDs that applied Maxwell's equation as a fundamental approach. However, previous publications claimed that the QDs had not been summarised computationally, which means the properties of the QDs have not been simulated [21, 23].

In this perspective, the idea of this study deals with a simple second generation of solar cells, the aSiSC model design combined with a part of the third generation of solar cells, a monolayer core/shell CdSe/ZnS spherical QD. The simulation used the FEM via COMSOL Multiphysics software. This study measures the light absorptions by aSiSC and aSiQDSC models and determines the advantages by applying CdSe/ZnS QD to the aSiSC model. Other than that, this study aimed to understand the physics concept behind the process and which influences the optical properties of

the QD with three different core sizes of CdSe/ZnS in the aSiQDSC model. Consequently, this model design is promising environment-friendly and high efficiency solar cells and will become a basic model for future research and support to future fabrication on the solar cell.

9.2 Simulation Models and Mathematical Approaches

The aSiSC model was adapted from [24]. Figure 9.1a shows design models with 9 nm width W and length L . The layers included were 10 nm aluminium (Al), 26 nm amorphous silicon (aSi), 5 nm indium tin oxide (ITO), and 5 nm air layer. Then, 9.0 nm thickness of CdSe/ZnS QD monolayer sandwiched in between ITO and aSi layers, known as the aSiQDSC model. The different CdSe sizes were modelled as shown in Fig. 9.1b, c, and d with 2.5 nm, 4.0 nm, and 7.0 nm of core diameter d_c , respectively. Table 9.1 shows the list of the properties of the materials from the previous researches used in this study.

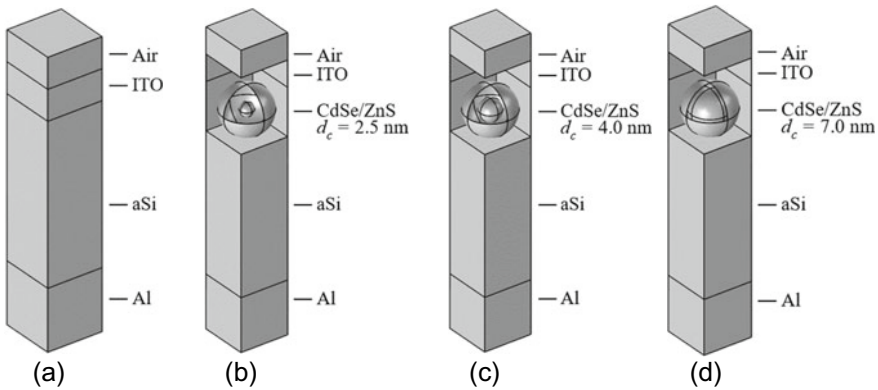


Fig. 9.1 Schematic diagram of the models **a** aSiSC, **b** aSiQDSC with $d_c = 2.5$ nm, **c** aSiQDSC with $d_c = 4.0$ nm, and **d** aSiQDSC with $d_c = 7.0$ nm

Table 9.1 The list of the properties of the materials from the previous researches used in this study

Material	Refractive index: real, n	Refractive index: imaginary, k	References
Al	0.69180	5.6354	[27]
aSi	4.3046	0.59238	[28]
CdSe	2.6187	0.3681	[29]
ZnS	2.387	0.037353	[30]
ITO	1.9016	0.0022439	[31]
Air	1.00027717	0	[32]

This simulation used the Finite Element Method to scrutinise the electric field profiles and light absorption by both aSiSC and aSiQDSC models. The mathematical models used in this study were Maxwell's equations, Fresnel's equation, and Beer-Lambert Law. The general equations of the electromagnetic wave (EMW) are given by Eqs. (9.1–9.4) [25] where Eq. (9.1) and (9.2) are steady state equations, Eqs. (9.3) and (9.4) are Faraday's law and Ampere's law with time-varying displacement current equations [26] where ρ is charge density and t is time.

$$\nabla \cdot \mathbf{B} = 0 \quad (9.1)$$

$$\nabla \cdot \mathbf{D} = \rho \quad (9.2)$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \quad (9.3)$$

$$\nabla \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t} + \mathbf{J} \quad (9.4)$$

The incident EMWs have interacted when propagating through the different materials with permittivity ε , permeability μ , and conductivity σ . The relationship between the electrical properties and electric displacement \mathbf{D} , current density \mathbf{J} , magnetic field \mathbf{H} , electric field \mathbf{E} , and magnetic induction \mathbf{B} can be represented as

$$\mathbf{D} = \varepsilon \mathbf{E} \quad (9.5)$$

$$\mathbf{B} = \mu \mathbf{H} \quad (9.6)$$

$$\mathbf{J} = \sigma \mathbf{E} \quad (9.7)$$

and the equations can be applied to Eqs. (9.3) and (9.4), yielding

$$\nabla \times \mathbf{E} = -\mu \frac{\partial \mathbf{H}}{\partial t} \quad (9.8)$$

$$\nabla \times \mathbf{H} = \varepsilon \frac{\partial \mathbf{E}}{\partial t} + \sigma \mathbf{E} \quad (9.9)$$

The curl vector of Maxwell's equation given by,

$$\nabla \times \nabla \times \mathbf{H} = \nabla(\nabla \cdot \mathbf{H}) - \nabla^2 \mathbf{H} = -\nabla^2 \mathbf{H} \quad (9.10)$$

But, in the free region case, the right side of the Eq. (9.10) equals zero. By using curl of \mathbf{E} and applying Eqs. (9.3) and (9.4) gave the EMW in the time domain

$$\begin{aligned}
\nabla \times \nabla \times \mathbf{E} &= -\nabla^2 \mathbf{E} \\
&= \mu \frac{\partial}{\partial t} (\nabla \times \mathbf{H}) \\
&= \mu \frac{\partial}{\partial t} \left(\frac{\partial \mathbf{D}}{\partial t} + \mathbf{J} \right) \\
&= \mu \frac{\partial}{\partial t} \left(\varepsilon \frac{\partial \mathbf{E}}{\partial t} + \sigma \mathbf{E} \right) \\
&= \mu \varepsilon \frac{\partial^2 \mathbf{E}}{\partial t^2} + \mu \sigma \frac{\partial \mathbf{E}}{\partial t}
\end{aligned} \tag{9.11}$$

The EMW equations applied in FEM by using Fourier transform to obtain the EMW equations in the frequency domain by derivative of $e^{i\omega t}$ to become

$$\frac{\nabla^2 \mathbf{E}}{\mu_r} - k^2 \left(\varepsilon_r - \frac{i\sigma}{\omega \varepsilon_0} \right) \mathbf{E} = 0 \tag{9.12}$$

where μ_r , ε_r , c , ω , σ , and k are relative permeability, relative permittivity, speed of light, angular frequency, electrical conductivity, and wavenumber, respectively [33].

Fresnel's equation is used in this simulation [34] as the EMW propagating in different materials and coupled with boundary conditions occurring reflection and refraction phenomena. The reflection R and transmission T coefficients of the Fresnel's equation for s - and p -polarisation are defined as

$$R_s = \frac{n_1 \cos \theta_i - n_2 \cos \theta_t}{n_1 \cos \theta_i + n_2 \cos \theta_t} \tag{9.13}$$

$$R_p = \frac{n_2 \cos \theta_i - n_1 \cos \theta_t}{n_1 \cos \theta_t + n_2 \cos \theta_i} \tag{9.14}$$

$$T_s = \frac{2n_1 \cos \theta_i}{n_1 \cos \theta_i + n_2 \cos \theta_t} \tag{9.15}$$

$$T_p = \frac{2n_1 \cos \theta_i}{n_1 \cos \theta_t + n_2 \cos \theta_i} \tag{9.16}$$

where R_s is TE reflection coefficient, R_p is TM reflection coefficient, T_s is TE transmission coefficient, and T_p is TM transmission coefficient, θ_i is the incident angle, θ_t transmitted angle, n_1 and n_2 is the first and second refractive index of the material, TE and TM are denoted as transverse electric and magnetic waves propagation.

The intensity of the incident light I_0 propagating onto the material and is scattered through with the models. The light propagates passing through the materials, and then transmitted light intensity I is observed lower than its original intensity. This is because the material properties attenuated the light intensity, and this is known as light absorption. Where the light passing through the materials being absorbed. The

Table 9.2 The semiconductor physical constants for CdSe, ZnS, and aSi [37]

Material	Mobility (cm ² V ⁻¹ s ⁻¹)		Effective mass		Electron affinity (eV)	Dielectric constant
	μ_e	μ_h	m_e/m_0	m_h/m_0		
aSi	1350	450	0.97	0.16	4.01	11.7
CdSe	600	9	0.13	0.45	4.95	10.0
ZnS	120	5	0.25	0.50	3.90	8.0

materials which have good light absorption gave lower transmitted light intensity than poor light absorption. The light absorption can be calculated by using Beer-Lambert law and described as [35, 36]

$$A = \log_{10} \left(\frac{I}{I_0} \right) \quad (9.17)$$

where A is absorption. The Eqs. (9.12–9.17) were used in the simulation to measure the electric field profiles and light absorption by the models.

The simulation then plotted the energy bandgap structures and the excited electron–hole concentrations for CdSe/ZnS QD and aSi by using its semiconductor physical constants from [37] as shown in Table 9.2. Both energy bandgap structures and the excited electron–hole concentrations were presented to relate and elaborate the electric field profiles within the models and the effect of the light absorption with different QD core sizes.

9.3 Results and Discussions

The model of aSiSC and aSiQDSC models were successfully designed and simulated. The light absorption results were shown graphically by the absorption over the wavelength for aSiSC and aSiQDSC with 2.5 nm, 4.0 nm, and 7.0 nm of core diameter of CdSe/ZnS QDs as shown in Fig. 9.2. Table 9.3 shows the light absorption peaks by aSiSC and aSiQDSC models with different QD core diameters.

Figure 9.2a shows the light absorption trends were higher in the UV region (200–380 nm). This is because the light energy in the UV region was more significant than the bandgap of the materials. Therefore, it can easily excite the electrons from the valence band (VB) to the conductive band (CB) of the materials. From 380 to 1000 nm, the light absorption started to climb and peaks at the middle of the visible region and then gradually decreased to 1000 nm, near the infrared region. The light absorption peaks in the visible region occurred as the light frequency resonating with the natural frequency of the materials, thus able to excite the electron from VB to CB. The materials were poorly absorbed in the near-infrared region due to low light energy to make electronic transitions. The light frequency is greater than the natural frequency of the materials, resulting in a lower transition of electrons [38].

Fig. 9.2 The light absorption curves **a** from 200 to 1000 nm and **b** light absorption peak region

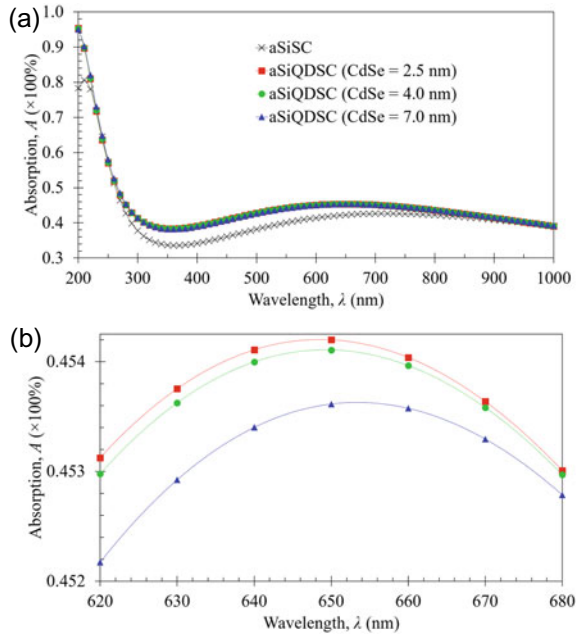


Table 9.3 The light absorption peaks by aSiSC and aSiQDSC models

Model	Core diameter, d_c (nm)	Light absorption peak, λ_m (nm)	Absorption, A (%)
aSiSC	0.0	720	42.654
aSiQDSC	2.5	646	45.376
	4.0	650	45.369
	7.0	657	45.330

Figure 9.2b showed the light absorption peaks by each aSiQDSC model. It has been observed the light absorption peak by the aSiSC model was occurred to the wider wavelength due to the absorption range of aSi was near the infrared region [39, 40]. Meanwhile, the light absorption peaks by aSiQDSC models had higher light absorption and occurred towards smaller wavelengths as the diameter of the QD core decreased. According to some experiments, the light absorption peaks by single CdSe/ZnS QDs were around 500–610 nm [40] and 513–562 nm [41]. Commonly, the absorption range by CdSe/ZnS QDs took place at shorter wavelengths even they are having larger sizes [40, 42]. Minor differences occurred by the aSiQDSC models as the massive presence of other materials in the models influenced the entire absorption range, which is dominantly by the aSi absorption range; meanwhile, the QD monolayer was small.

From the results obtained, the presence of QD layer in aSiQDSC models with 2.5 nm, 4.0 nm, and 7.0 nm of QDs core diameter gave better light absorptions by 6.00%, 5.98%, and 5.90% of increment, respectively, as compared to the aSiSC model. Three factors played a role in the increment and peaks of light absorption by the aSiQDSC models: multi-exciton generation (MEG) occurred within the QD, the classification of the QDs, and the difference in core diameter of the QDs which gave different energy bandgap due to the quantum confinement effect.

In a bulk semiconductor, the incident light with higher energy than the semiconductor bandgap able to excite an electron from the VB to the higher level in the CB and the electron known as a hot carrier electron. The hot carrier electron then undergoes many nonradiative relaxations continuous transitions, which is thermalisation, before going to the lower level of CB. Meanwhile, the quantum confinement effect of the QD results in a large and discrete bandgap level in both VB and CB. The hot carrier electron undergoes light emission as going down to the lower level in CB. The light emission has higher energy than the bandgap, thus being absorbed by another electron at the VB [13]. The phenomenon that occurred inside the QD was known as MEG. A single incident photon can generate up to two or three excitons, producing more photons when the recombination occurs. This process has many photons that are totally from both incident light and by the QD and gave benefits to the aSi in aSiQDSC models compared to the aSiSC model, which only received photons totally from the incident light.

In terms of QD classification, CdSe/ZnS was classified as type-I core/shell QD, which comprises a broader bandgap of shell and a small bandgap of the core. Type-I QD improves the light absorption in the visible light region when applied to the solar cell. Therefore, it influences the blue shifting of light absorption by the aSiQDSC model compared to the aSiSC model. In addition, both carriers (excited electrons and the holes) were confined inside the core, thus emitting light energy core dependent on the aSi [43]. Figure 9.3 shows the electric field profiles within the entire model for each aSiQDSC model at their respective light absorption peaks. The results from Fig. 9.3 clearly showed higher electric fields occurred in aSi by 2.5 nm diameter core compared to the model with 4.00 nm followed by 7.00 nm. The aSi within the aSiQDSC models absorbed photons from the incident photon and MEG from the core of QD.

The QDs with different core diameters and shell thickness gave different light absorption and light absorption peaks. This is because the quantum confinement effect changed the semiconductor materials' energy bandgap, making the energy bandgap of the QDs tuneable [43]. Table 9.4 shows the value of the energy bandgap for CdSe/ZnS with different core sizes. Figure 9.4 shows the energy bandgap structures and the concentration of electrons and holes for CdSe/ZnS and aSi in each aSiQDSC model.

From Table 9.4 and Fig. 9.4a, c, and d, as the core diameter smaller, the energy bandgap wider, thus resulting in higher light absorption and occurred toward the smaller wavelength (blue-shifting). The phenomena happened inside the core itself, where both electrons and holes carriers were confined in the core. Therefore, the recombination that occurred in the core gave the light emission depends on the

Fig. 9.3 The electric field profiles within the entire model **a** aSiQDSC with $d_c = 2.5$ nm, **b** aSiQDSC with $d_c = 4.0$ nm, and **c** aSiQDSC with $d_c = 7.0$ nm

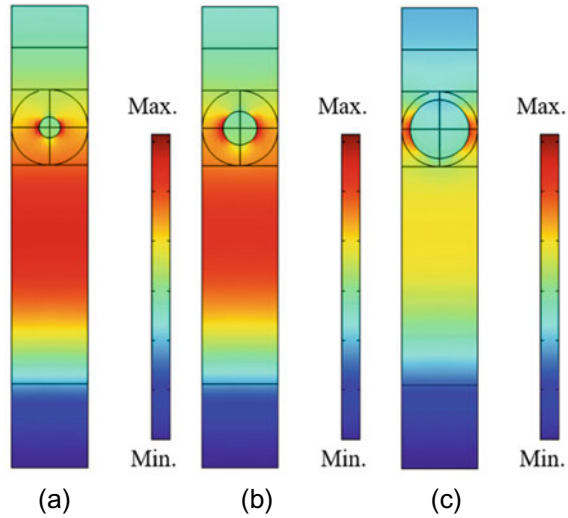


Table 9.4 The core energy bandgap, optical absorption and its light peaks of CdSe/ZnS for each core diameter

Core diameter, d_c (nm)	Core energy bandgap, E_{gc} (eV)	Light absorption peak, λ_m (nm)	Absorption, A (%)
2.5	3.88	646	45.376
4.0	2.50	650	45.369
7.0	1.93	657	45.330

energy bandgap of the core. The small core diameter has a wider energy bandgap, the electrons excited as absorbing higher photon energy—the emission by MEG and the recombination also emitting high. Thus, the aSi absorbed high energy of the photon, and the peak occurred at a smaller wavelength. On the other hand, a large core diameter has a smaller energy bandgap; the recombination emitted with low energy photon and peaks occurred towards the larger wavelength. This also explained the electric fields in Fig. 9.3.

Figure 9.4b, d, and f show the concentration of the excited carriers of electrons and holes for each aSiQDSC model. The excited electrons and holes were dominantly conquered in the core region, where the recombination occurred in the core, as mentioned from [43]. As the core diameter increased, the value of exciting electrons decreased thus, low recombination occurred. This also can be related to electric field profiles within the core in Fig. 9.3. The electric fields inside the core were lower than the shell, and this is because a bunch of excited electrons and holes recombined within the core at the wavelength peak, thus lower electric fields within the core. By comparing with three different core diameters, the electric field profiles within 2.5 nm core were higher than 4.0 nm and followed by 7.0 nm.

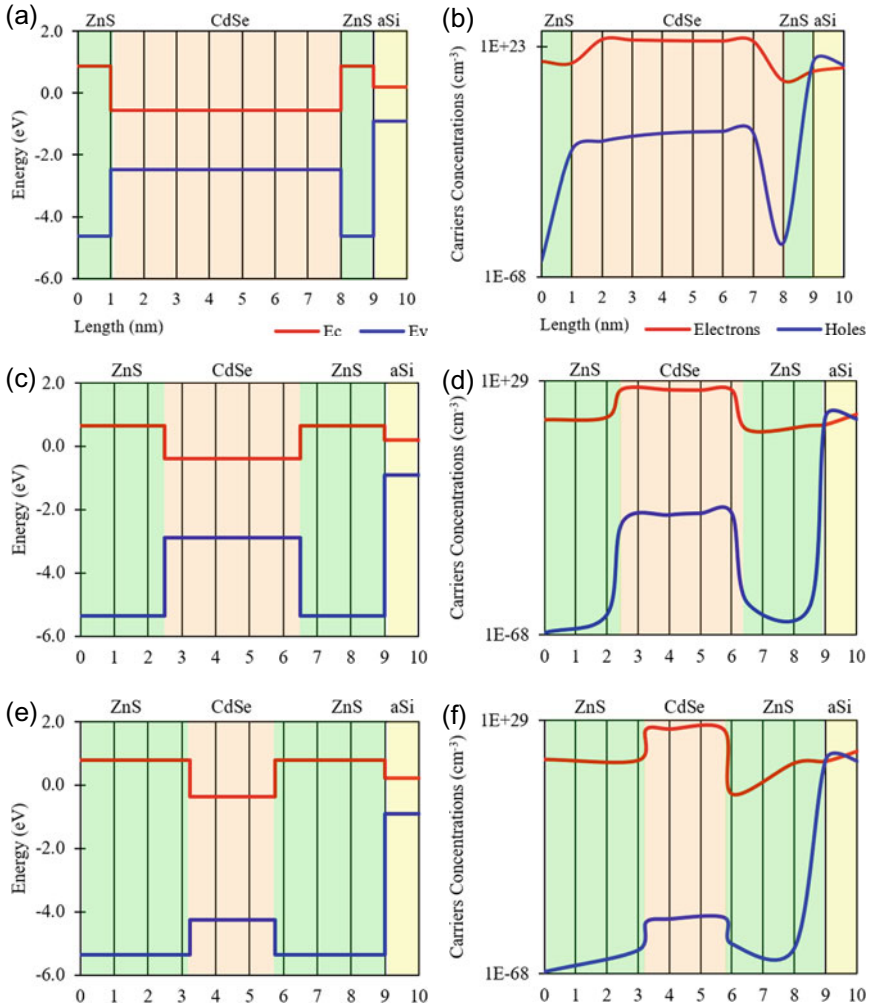


Fig. 9.4 The energy band gap structures and the excited electrons and holes concentrations for CdSe/ZnS and aSi with **a, b** $d_c = 2.5$ nm, **c, d** $d_c = 4.0$ nm, and **e, f** $d_c = 7.0$ nm

These simulated results have proved the statements from [2, 7, 17, 18] and the experimented results from [44], where the different sizes of the spherical CdSe/ZnS QDs appeared in the various peaks of the light absorptions and the peaks of the light absorptions occurred at the different wavelength. All things considered in these simulated results, the presence of QD monolayer applied on the aSiSC models increase the light absorption and the larger the core side of CdSe/ZnS QD, the lower the peak of the light absorption and the light of absorption occurred at the larger wavelength and vice versa.

9.4 Conclusion

In summary, the aSiSC and aSiQDSC were successfully designed via COMSOL Multiphysics and simulated using the Finite Element Method (FEM) to measure the light absorptions of the models. The results obtained, such as electric fields, energy bandgap of the quantum dots, and the carriers' concentration, were explained to determine the factor that influences the light absorption by different core diameters of the CdSe/ZnS QDs. The light absorptions were analysed in the range of 200 nm to 1000 nm of the wavelength. The light absorption peaks by aSiSC model were occurred at larger wavelengths due to the aSi in the models. The presence of QD in the models increased the light absorption. This is because the multi-exciton generations occurred within the QDs. The multiple photons can be generated and being absorbed by aSi layer. The absorption peaks that occurred were shifted to the shorter wavelengths when the core diameter of the QD decreased. This happened because the quantum confinement effect of the QD was able to tune the energy bandgap of the QDs. Therefore, the solar spectrum can be absorbed in the entire visible region with the presence of the QDs. The type of QDs also important as the material having its specific absorption ranges, which also influenced the light absorption by the models. For future study, some nanostructures such as nanocavity and nanoantenna, which enhanced the light trapping, will be applied on the aSiQDSC to measure light absorption effectiveness by the models.

Acknowledgements This study was supported by the Ministry of Higher Education (MOHE) through Fundamental Research Grant Scheme (FRGS) (FRGS/1/2019/STG02/UTHM/02/1) and Universiti Tun Hussein Onn Malaysia (UTHM) through vot K173.

References

1. Y.I. Abdulkarim, L. Deng, F.F. Muhammad, L. He, Enhanced light absorption in the organic thin films by coating cross-shaped metamaterial resonators onto the active layers. *Results Phys.* **13**, 102338 (2019)
2. Z. Ahmad, M.A. Najeeb, R.A. Shakoor, S.A. Al-Muhtaseb, F. Touati, Limits and possible solutions in quantum dot organic solar cells. *Renew. Sustain. Energy Rev.* **82**, 1551–2156 (2018)
3. S. Emin, S.P. Singh, L. Han, N. Satoh, A. Islam, Colloidal quantum dot solar cells. *J. Sol. Energy* **85**, 1264–1282 (2011)
4. S. Felicetti, *Ultrastrong light-matter interaction in quantum technologies* (Universidad del Pais Vasco-Euskal Herriko Unibertsitatea, 2015)
5. P.Tighineanu, *Electric and magnetic interaction between quantum dots and light* (University of Copenhagen, 2015)
6. A. Delgado, M. Korkusinski, P. Hawrylak, Theory of atomic scale quantum dots in silicon: Dangling bond quantum dots on silicon surface. *Solid State Commun.* **305**, 113752 (2020)
7. K.E. Jasim, *Solar Cells: New Approaches and Reviews*, ed. by L.A. Kosyachenko, pp. 303–327 (IntechOpen, London, 2015)

8. T.K. Nideep, M. Ramya, M. Kailasnath, The influence of ZnS buffer layer on the size dependent efficiency of CdTe quantum dot sensitized solar cell. *Superlattices Microstruct.* **130**, 175–181 (2019)
9. M.J. Schütz, *Quantum Dots for Quantum Information Processing: Controlling and Using their Environment* (Ludwig-Maximilian University, 2015)
10. M. Slamet, V. Sahni, Electron correlations in an excited state of a Quantum Dot in a uniform magnetic field. *Comput. Theor. Chem.* **1114**, 125–139 (2017)
11. S.M. Ali, S. M. Ramay, M. H. Aziz, N. Ur-Rehman, M.S. AlGarawi, S.S. AlGhamd, A. Mahmood, T.S. Alkhurajji, S. Atiq, Efficiency enhancement of perovskite solar cells by incorporation of CdS quantum dot through fast electron injection. *Org. Electron.* **62**, 21–25 (2018)
12. T.M.W.J. Bandara, J.L. Ratnasekera, Polymer electrolytes for quantum dot-sensitized solar cells (QDSSCs) and challenges, in *Polymer Electrolytes: Characterization Techniques and Energy Applications*. ed. by T. Winie, A.K. Arof, S. Thomas (Wiley-VCH, Weinheim, 2020), pp. 299–337
13. K. Surana, R.M. Mehra, B. Bhattacharya, Quantum dot solar cells with size tuned CdSe QDs exhibiting 1.51 V. *Mater. Today Proc.* **5**, 9108–9113 (2018)
14. P. Venkatachalam, S. Rajalakshmi, Performance of perovskite and quantum dot sensitized solar cell based on ZnO photoanode structure. *Mater. Today: Proc.* **22**, 400–403 (2020)
15. L. Hu, X. Geng, S. Singh, J. Shi, Y. Hu, S. Li, X. Guan, T. He, X. Li, Z. Cheng, R. Patterson, S. Huang, T. Wu, Synergistic effect of electron transport layer and colloidal quantum dot solid enable PbSe quantum dot solar cell achieving over 10% efficiency. *Nano Energy* **64**, 103922 (2019)
16. M. Xing, Y. Zhang, Q. Shen, R. Wang, Temperature dependent photovoltaic performance of TiO₂/PbS heterojunction quantum dot solar cells. *Sol. Energy* **195**, 1–5 (2020)
17. V.G. Reshma, P.V. Mohanan, Quantum dots: Applications and safety consequences. *J. Lumin* **205**, 287–298 (2019)
18. K.E. Jasim, *Third-Order Nonlinear Optical Properties of Quantum Dots*, ed. By C. Luigi, pp. 1–19 (IntechOpen, London, 2019)
19. D.B. Hayrapetyan, D.A. Baghdasaryan, E.M. Kazaryan, S.I. Pokutnyi, H.A. Sarkisyan, Exciton states and optical absorption in core/shell/shell spherical quantum dot. *Chem. Phys.* **506**, 26–30 (2018)
20. Y. Hong, Y. Wu, S. Wu, X. Wang, J. Zhang, Overview of computational simulations in quantum dots. *Isr. J. Chem.* **59**, 661–672 (2019)
21. W.K. Cartar, *Theory and Modelling of Light-Matter Interactions in Photonics Crystal Cavity Systems Coupled to Quantum Dot Ensembles* (Queen's University, 2017)
22. I. Richter, P. Kwiecien, J. Ctyroky, Advanced photonic and plasmonic waveguide nanostructures analyzed with Fourier modal methods, in 2013 15th International Conference on Transparent Optical Networks, pp. 1–7 (Institute of Electronic and Electronics Engineers, Spain, 2013)
23. Z. Hu, L. Jensen, A discrete interaction model/quantum mechanical method for simulating plasmon-enhanced two-photon absorption. *J. Chem. Theory Comput.* **14**, 5896–5903 (2018)
24. Z. Zhu, J. Sun, Z. Li, X. Yu, J. Zhao, H. Dai, Tunable light absorption of Ag nanoparticles@ carbon sphere by the position control of Ag nanoparticles. *Optik* **179**, 831–836 (2019)
25. D. Felbacq, *Advanced Numerical Techniques for Photonic Crystals* (Morgan & Claypool Publishers, 2016)
26. H.J. Pain, *The Physics of Vibrations and Waves* (John Wiley & Sons, Ltd, England, 2021)
27. F. Cheng, P.H. Su, J. Choi, S. Gwo, X. Li, C.K. Shih, Epitaxial growth of atomically smooth aluminum on silicon and its intrinsic optical properties. *ACS Nano* **10**, 9852–9860 (2016)
28. D.T. Pierce, W.E. Spicer, Electronic structure of amorphous Si from photoemission and optical studies. *Phys. Rev. B* **5**, 3017 (1972)
29. S. Ninomiya, S. Adachi, Optical properties of cubic and hexagonal CdSe. *Jpn. J. Appl. Phys.* **78**, 4681–4689 (1995)
30. S. Ozaki, S. Adachi, Optical constants of cubic ZnS. *Jpn. J. Appl. Phys.* **32**, 5008 (1993)

31. P.E. Ciddor, Refractive index of air: new equations for the visible and near infrared. *Appl. Opt.* **35**, 1566–1573 (1996)
32. R.J. Moerland, J.P. Hoogenboom, Subnanometer-accuracy optical distance ruler based on fluorescence quenching by transparent conductors. *Optica* **3**, 112–117 (2016)
33. V. Giannini, N.P. Hylton, X. Li, K.H. Lee, N.J. Ekins-Daukes, S.A. Maier, Solar cells with a multi-functional plasmonic light concentration layer, in *2011 37th IEEE Photovoltaic Specialists Conference*, pp. 000864–000865 (IEEE, 2011)
34. O. Lehtikangas, T. Tarvainen, A.D. Kim, S.R. Arridge, Finite element approximation of the radiative transport equation in a medium with piece-wise constant refractive index. *J. Comput. Phys.* **282**, 345–359 (2015)
35. M.H. Muhammad, M.F.O. Hameed, S.S.A. Obayya, Absorption Enhancement in Hexagonal Plasmonic Solar Cell, in *Numerical Simulation of Optoelectronic Devices*, pp. 63–64 (Institute of Electronic and Electronics Engineers, Spain, 2014)
36. M. Hu, J. Chen, Z.Y. Li, L. Au, G.V. Hartland, M. Marquez, Y. Xia, Gold nanostructures: engineering their plasmonic properties for biomedical applications. *Chem. Soc. Rev.* **35**, 1084–1094 (2006)
37. S. Kar, *Advanced Microelectronics* (Springer Series, Berlin, 2013)
38. A. Shah, *McEvoy's Handbook of Photovoltaics* (EPFL Press, Lausanne, 2018)
39. T.N. Truong, D. Yan, C. Samundsett, R. Basnet, M. Tebyetekerwa, L. Li, F. Kremer, A. Cuevas, D. Macdonald, H.T. Nguyen, Hydrogenation of phosphorus-doped polycrystalline silicon films for passivating contact solar cells. *ACS Appl. Mater. Interfaces* **11**, 5554–5560 (2019)
40. S. Prasad, H.S. AlHesseny, M.S. AlSalhi, D. Devaraj, V. Masilamai, A high power, frequency tunable colloidal quantum dot (CdSe/ZnS) laser. *Nanomaterials* **7**, 29 (2017)
41. B.M. Saidzhonov, V.F. Kozlovsky, V.B. Zaytsev, R.B. Vasiliev, Ultrathin CdSe/CdS and CdSe/ZnS core-shell nanoplatelets: The impact of the shell material on the structure and optical properties. *J. Lumin.* **209**, 170–178 (2019)
42. R. Hamood, M.S. El-sadek, A. Gadalla, Facile synthesis, structural, electrical and dielectric properties of CdSe/CdS core-shell quantum dots. *Vacuum* **157**, 291–298 (2017)
43. A.R. AbouElhamd, K.A. Al-Sallal, A. Hassan, Review of core/shell quantum dots technology integrated into building's glazing. *Energies* **12**, 1058 (2019)
44. V. Krivenkov, P. Samokhvalov, M. Zvaigzne, I. Martynov, A. Chistyakov, I. Nabiev, Ligand-mediated photobrightening and photodarkening of CdSe/ZnS quantum dot ensembles. *J. Phys. Chem. C.* **122**, 15761–15771 (2018)

Chapter 10

Electronic Structure and Optical Properties of Lead-Free Double Perovskite BaMgLaBiO₆: A DFT Study



Junaid Munir, Muhammad Jamil, Quratul Ain, and Kaneez Fatima

Abstract Structural, electronic and optical properties of double perovskite BaMgLaBiO₆ are theoretically investigated with density functional theory (DFT) implemented in WIEN2k code. The structure is optimized to achieve minimum energy at the ground state and optimized lattice parameters and bulk moduli are obtained. The formation energy is also calculated which confirms the possibility to synthesize the studied compound. Generalized gradient approximation (GGA) and modified Becke-Johnson (mBJ) exchange potentials are used to calculate the physical properties. A semiconductor behavior is observed with respective potentials. A band-gap of 2.7 eV and 3.8 eV is calculated with GGA and mBJ potentials respectively. The electron density plots show the both covalent and ionic bonding between the atoms in the compound. Total density of states is also calculated which shows a good agreement with the band structure. The optical parameters such as real and imaginary parts of dielectric function, refractive index, extinction coefficient, optical conductivity, absorption coefficient, and energy loss function are also calculated. A good optical conductivity is achieved in the selected energy range. On the basis of calculated properties, this compound is recommended as a potential candidate in LED's, lasers, and power switching applications.

10.1 Introduction

The double perovskite oxides have attracted much attention because of their chemical flexibility and extensiveness of configurationally space spanning [1]. In the past few years, double perovskites have been studied extensively due to their extraordinary properties and promising technological applications [2–4]. They have obtained fundamental interest because of numerous interesting physical properties including

J. Munir (✉) · M. Jamil · K. Fatima
Department of Physics, Riphah International University, Lahore, Pakistan
e-mail: junaid.munir@riphah.edu.pk

Q. Ain
Department of Physics, University of Management and Technology, Lahore, Pakistan

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2022
A. B. Mustapha et al. (eds.), *Proceedings of the 7th International Conference on the Applications of Science and Mathematics 2021*, Springer Proceedings in Physics 273, https://doi.org/10.1007/978-981-16-8903-1_10

spin-polarized half-metallic electrical conductivity, their unique electronic structure ranging from metallic to insulator [5–7], and superconductivity [8]. The perovskite structure with general formula ABX_3 (Space group $Pm-3_m$) with cubic symmetry is formed by connecting BX_6 octahedral through corners (Where A and B are cations). In the original perovskite, the charge on A and B cation may vary. Double perovskites are also the significant members of this family with different composition, structure and physical properties than the original perovskites. Double perovskites (general formula = $A_2BB'X_6$) are basically derived from perovskite have advanced technological applications in the field of magneto-dielectric, magneto-optics, half metallic and spintronics [9, 10]. In double perovskites, A is a monovalent cation which is larger than monovalent B cation and halide ion X acts as potential alternatives [11]. The double perovskite halides have large carrier lifetimes in comparison with lead-based perovskites [12] and shows high dispersed conduction and valence bands [13]. Recently, Ba_2ZnOsO_6 is reported with cubic structure having $Fm-3_m$ (225) space group [14]. A large number of perovskites including double perovskites have been investigated for their physical properties [15–17]. Yun-Ping et al. [18] reported A_2CrMO_6 for half metallic properties based on first principle calculation. Mostly reported double perovskites have indirect band-gap in the range of 2–5 eV. It limits the utilization of solar spectrum effectively [13]. The degradation of $Cs_2AgBiBr_6$ is observed on exposure to light and ambient air [13]. The composition space for merging different cations A and B by using monovalent halides restricts to achieve the required properties [19]. Recently, optoelectronic magnetic and thermodynamic properties of double perovskite oxide have been studied [20–22].

To achieve a stable structure of perovskite and address the above-mentioned challenges, a new class of double perovskites with a general stoichiometry $A'A''B'BiO_6$ containing Bi^{3+} cations in the 3D octahedral framework is introduced [4]. The structures related to this class are widely studied and show good properties for practical devices [23–25]. In the present work, we have studied the lead-free $BaLaMgBiO_6$ double perovskite under the framework of density functional theory (DFT) using wien2k code. According to the author's best knowledge, so far no DFT-based study is done on the said compound to explore its properties. We will calculate the structural, electronic, and optical properties of $BaLaMgBiO_6$ compound to investigate its response for optoelectronic applications.

10.2 Computational Details

Electronic structure and geometry optimization is done with full potential linearized augmented plane wave (FPLAPW) method implemented in WIEN2k code [26]. Kohn–Sham equations are calculated with FPLAPW with the help of self-consistent method [27]. In Kohn–Sham DFT, the exchange correlation energy ($E_{XC} = E_X + E_C$) as a functional of electron spin densities must be approximated. For slow varying densities, the famous functional have appropriate form: the generalized gradient

approximation (GGA). Perdew, Burke and Ernzerhof Generalized Gradient Approximation (PBE-GGA) [28] is used to solve the electron exchange and correlation potential. However, the utilization of GGA strongly depends upon the properties of studied materials. The results achieved with GGA for the band-gap of insulators and semiconductors are severely not in good agreement with the experimental results, e.g., it mostly underestimated the band-gap or sometimes even absent [29]. To solve this problem and achieve more accuracy in band-gap, modified Becke-Johnson exchange potential (mBJ) [30] is used. The total Hamiltonian of the system is calculated accurately with mBJ approximation because it uses both GGA exchange and correlation potentials. As a result, the band-gap accuracy comparable with experimental result is significantly improved [30]. GGA correlation with mBJ potential is used in our calculation for more accurate band-gap. The values for selected muffin tin radii (a.u) of different atoms are Ba = 2.50, La = 2.50, Mg = 1.83, Bi = 2.01 and O = 1.69. The integration is done with tetrahedral method using 1000 k-points in the complete Brillouin zone. The plane wave expansion cutoff is set at RMT K_{\max} = 7.0 to carry the expansion of wave function, where RMT = atomic sphere radius and K_{\max} = plane wave cutoff. The electron potential and electron density inside the muffin tin spheres are presented with spherical harmonics [31] is set to l_{\max} = 10. The self-consistency convergence criterion for energy is achieved up to 10⁻⁵ Ry.

10.3 Results and Discussion

10.3.1 Structural and Electronic Properties

BaMgLaBiO₆ has a cubic double perovskite structure (general formula $A'A''B'B''X_6$) with a space group of F-43 m. It is the extension of perovskite structure ABX₃. The oxygen atoms (O₂) are bounded with Ba⁺² to form cuboctahedra BaO₁₂ and it share corners with other equivalent BaO₁₂. Similarly, the oxygen atoms (O₂) are bounded with Mg⁺², La⁺³, and Bi⁺⁵ to form cuboctahedra MgO₁₂, LaO₁₂, and BiO₁₂ that also share corners with other cuboctahedra. The tilt angles of corner-sharing octahedra are 0°. The bond lengths for all Ba–O and Mg–O, La–O and Bi–O are 2.93 Å, 2.05 Å, 2.93 Å and 2.10 Å respectively. In a linear distorted geometry, O₂ is bounded to two Ba⁺², one Mg⁺², two La⁺³ and one Bi⁺⁵ atom. Figure 10.1a shows the crystallographic unit cell of double perovskite BaMgLaBiO₆.

The unit cell of BaMgLaBiO₆ has been optimized for the ground state stable atomic configuration and to achieve optimized lattice constants. The Murnaghan's equation of state is used to calculate the ground-state lattice parameters.

$$E(V) = E_0 + \left[\frac{\left(\frac{V_0}{V}\right)^{B_0}}{B_0 - 1} + 1 \right] - \frac{BV_0}{B_0 - 1} \quad (10.1)$$

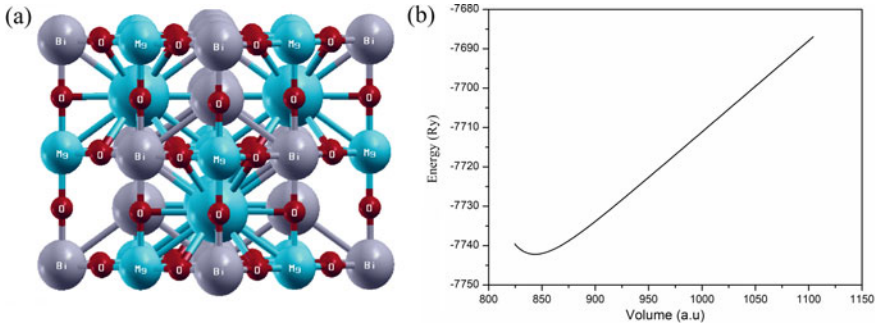


Fig. 10.1 Crystal structure (a) and energy-volume curve (b) of perovskite BaMgLaBiO₆

where E_0 is the minimum energy and V_0 is the volume at $T = 0$ k. the bulk modulus is B and pressure derivative of B is B_0 at the equilibrium volume. The optimization curve is presented in Fig. 10.1b. On the parabolic curve, the obtained minimum energy value shows the ground state energy and the corresponding volume is the ground state volume.

In addition, the formation energy is also calculated to check the possibility to synthesize this compound. The formation energy is basically the difference between the energies of the stable phases of the elements and the total energy of the compound.

$$E_{FE} = E_{BaMgLaBiO_6} - E_{Ba} - E_{Mg} - E_{La} - E_{Bi} - 6E_O \quad (10.2)$$

It can be seen from the calculated formation that the studied compound can be synthesized. The calculated lattice parameters such as bulk modulus B (GPa), pressure derivative of B and ground state energy E_0 (Ry) are presented in the Table 10.1.

The band structure is the key factor in describing the electronic behavior of any compound. Figure 10.2 shows the band structure of BaMgLaBiO₆ perovskite calculated with generalized gradient approximation (GGA) and modified Becke-Johnson (mBJ) potentials along the high symmetry points of 1st Brillouin zone.

Conduction band is observed above the Fermi level (0 eV) and valence band below it. The energy band-gap lies between the maxima of valence and minima of conduction band. A band-gap is observed and it increases from GGA to mBJ potential which confirm a semiconductor nature of this perovskite. The calculated values of band-gap are 2.7 eV and 3.8 eV with GGA and mBJ potential respectively. The band-gap nature can be further investigated with density of states (DOS) plots. The understanding of atomic/orbital contribution towards the band structure can be

Table 10.1 Optimized parameters of perovskite BaMgLaBiO₆

Lattice parameters $a(A^\circ)$	Bulk modulus B (GPa)	B'	$E_0(Ry)$	Formation energy	
BaMgLaBiO ₆	7.91	153.2	4.84	-7742.214	-4.11552

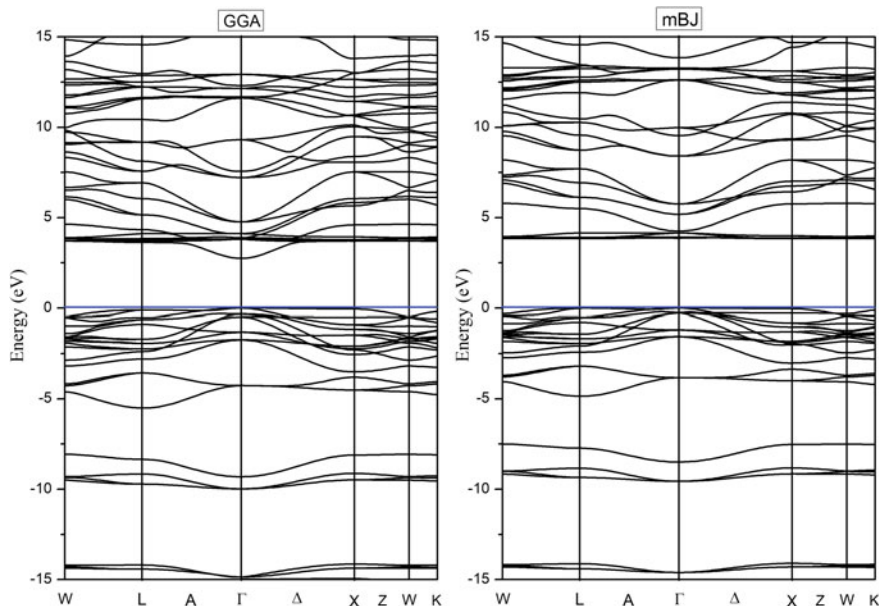


Fig. 10.2 Calculated band-structure of perovskite BaMgLaBiO_6 with GGA and mBJ potentials

achieved with DOS. Figure 10.3 show the total density of states as a function of energy. The selected energy range is from -15 eV to 15 eV. The conduction band is represented with positive energy range and the negative energy range corresponds to the valence band. A clear band-gap is observed near the Fermi level in the plot of total density of states with both potentials. The contribution of different states above

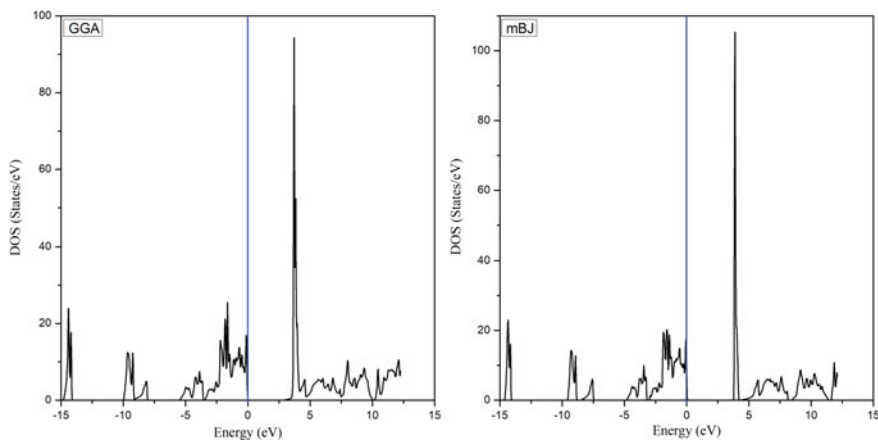


Fig. 10.3 Total density of states (TDOS) of perovskite BaMgLaBiO_6 with GGA and mBJ potentials

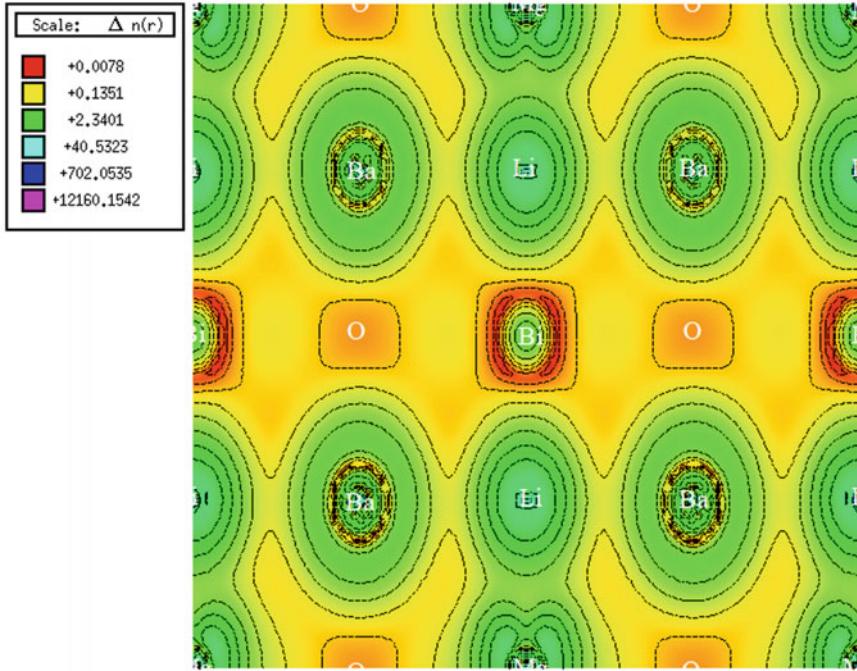


Fig. 10.4 Electron density along (110) plane of perovskite BaMgLaBiO₆

Fermi level in conduction band (CB) and below Fermi level in valence band (VB) start from the energy band-gap. The Ba-d and La-d states show their main contribution in the conduction and valence band. The s and p states of Mg and Bi have contribution in valence and conduction bands while the O-p state shows its major contribution in valence band. The contribution of other states for all atoms is negligibly small.

Electron density plots reveal the bonding nature between the atoms in the compound. The electron density along (110) plane for BaMgLaBiO₆ is shown in Fig. 10.4. A weak covalent bonding is observed between the Ba, and La atoms. Oxygen (O) atoms show the ionic bonding nature along with the Bi atoms. A strong bonding can be seen between the Bi and Mg atoms.

10.3.2 Optical Properties

The optical response of the compound is represented with dielectric function at photon's energies ($E = hc/\lambda$). Figure 10.4a shows the real $\epsilon_1(\omega)$ part of dielectric function of BaMgLaBiO₆ with GGA and mBJ potentials. The selected energy range is 0 to 14 eV. The static frequency limit $\epsilon_1(0)$ are 4.68 and 4.10 with GGA and mBJ potentials respectively. As the band-gap increases from GGA to mBJ potential

(Fig. 10.2), $\epsilon_1(0)$ values decreases respectively which shows an inverse relation between static frequency and band-gap as depicted in Penn model [32]. Beyond the static point, $\epsilon_1(\omega)$ start increasing and reaches a maximum value of 6.28 at 3.71 eV and 4.58 at 3.86 eV with GGA and mBJ respectively. It can be seen from the main characteristic peaks that the compound falls in the visible range of energy in electromagnetic spectrum. After the peaks, the plot observed a decrease and become negative at certain range of energy. The negative values show complete attenuation of light and depict metallic behavior [33]. Figure 10.4 represent the imaginary $\epsilon_2(\omega)$ part of dielectric function which shows the band-gap and absorptive behavior of the material. The imaginary part is also the representative of inter-band transitions. The $\epsilon_2(\omega)$ has two sharp peaks at 4.7 and 6.2 eV with GGA and 4.0 and 4.5 eV with mBJ potential. Depending upon the information collected from both $\epsilon_1(\omega)$ and $\epsilon_2(\omega)$ the calculation of essential optical parameters such as absorption coefficient, extension coefficient, refractive index and optical conductivity become permissible.

Refractive index ($n = c/v$) is the dimensionless quantity which describes the behavior of light travelling through the given medium. The refractive index $n(\omega)$ and extinction coefficients $k(\omega)$ both relates the interaction of light with material. The refractive index $n(\omega)$ and extinction coefficient $k(\omega)$ as a function of energy are calculated with respective potentials and plotted in Fig. 10.5. Figure 10.5a shows the behavior of $n(\omega)$ which follow the same pattern as imaginary part of dielectric function. The static frequency $n(0)$ values are 2.01 and 1.75 with GGA and mBJ

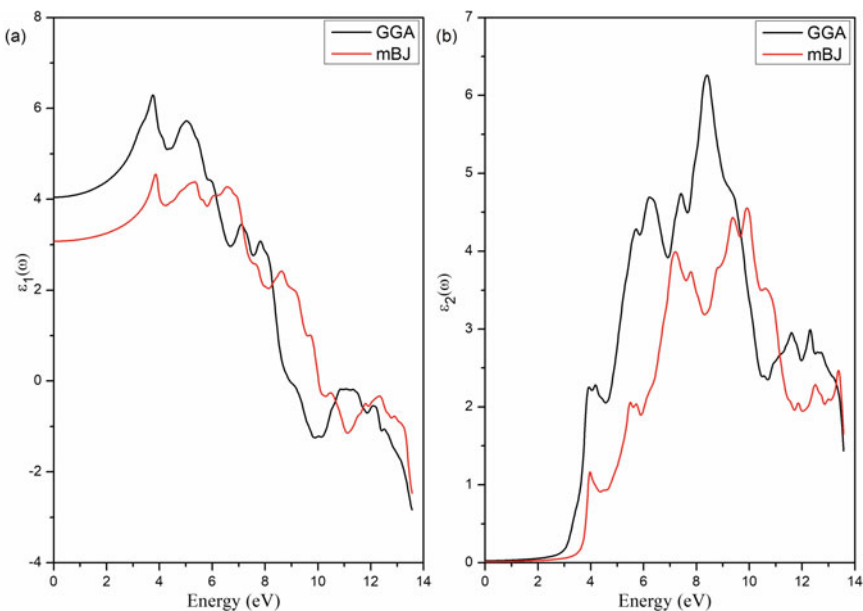


Fig. 10.5 The real $\epsilon_1(\omega)$ (a) and imaginary real $\epsilon_1(\omega)$ (b) parts of dielectric function of BaMgLaBiO₆ with GGA and mBJ potentials

potentials, respectively. It can be seen from band structure (Fig. 10.2) that the band-gap is larger for mBJ potential, while $n(0)$ (Fig. 10.5a) is smaller for mBJ potential which confirms an inverse relation between them. Further increase in energy causes to increase the value of $n(\omega)$ and maximum values obtained at 3.75 eV (2.51) and 3.83 eV (2.15) with GGA and mBJ respectively. There are also other peaks but a gradual decrease is observed with further increase in energy.

The extinction coefficient $k(\omega)$ is basically the imaginary part of refractive index $n(\omega)$ and plotted in Fig. 10.5b. It shows the absorbed radiation in an analogous way. The threshold energies of $k(\omega)$ are 2.1 eV and 2.7 eV with both potentials. The $k(\omega)$ values show an increase with increase in energy and maximum values observed at 8.5 and 9.6 eV for GGA and 10.1 and 11.0 eV for mBJ potential. The high peaks indicate the high absorption in the perovskite.

The optical conductivity spectra is calculated and plotted in Fig. 10.6a. It is revealed from the figure that optical conductivity starts from 2.7 eV and 3.8 eV that is the band gap calculated with with GGA and mBJ potential respectively. After the band-gap values, the optical conductivity gradually increases and achieves the peak values at 8.4 eV and 9.9 eV with respective potentials. The optical conductivity spectra start decreasing after the peak values with multiple minor peaks.

The relationship between the energies of incident photon and their per unit length absorption is illustrated with absorption coefficient $\alpha(\omega)$ spectrum as shown in Fig. 10.6b. A non-linear increase in $\alpha(\omega)$ as the function of energy. The abrupt

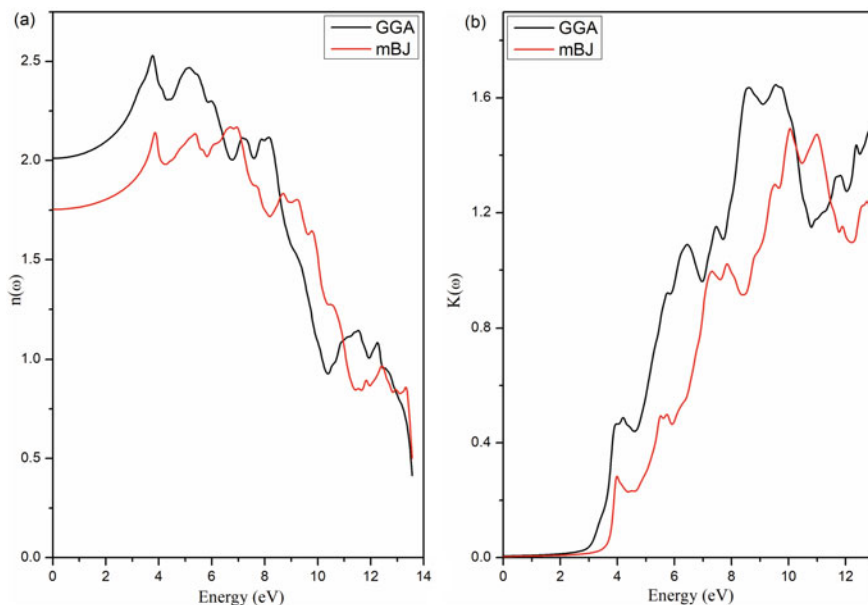


Fig. 10.6 Refractive index (a) and extinction coefficient (b) of BaMgLaBiO₆ calculated with GGA and mBJ potentials

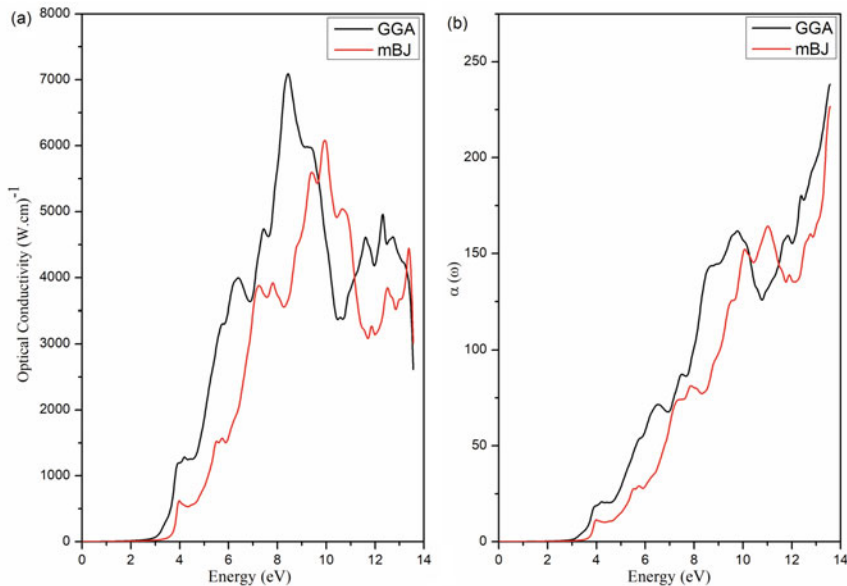


Fig. 10.7 The optical conductivity (a) and absorption coefficient (b) of BaMgLaBiO₆ calculated with GGA and mBJ potentials

increase in $\alpha(\omega)$ correspond to the highest absorption of incident radiations. The photon absorption is the property of semiconductors. The absorbed photon causes to excite and shifting the electrons from valence band to conduction band.

The energy loss function represents the loss in energy of moving electrons with high velocity by with the collision of other electrons while moving through the material.

Figure 10.7 shows the energy loss as a function of energy in BaMgLaBiO₆ perovskite. Several peaks are observed for energy loss function. The maximum values are obtained at 10.7 eV and 12.1 eV with GGA and mBJ potentials, respectively (Fig. 10.8).

10.4 Conclusion

The electronic structure and optical properties of double perovskite BaMgLaBiO₆ are studied in the presented work. The full-potential linearized augmented plane wave (FP-LAPW) method is employed under DFT framework via WIEN2k code. A semiconductor behavior revealed from the electronic properties with GGA and mBJ potentials. The density of states also gives the same results as band structure. The electron density is also studied and presented in the contour plots. Different optical parameters are investigated which shows the potential of the compound for

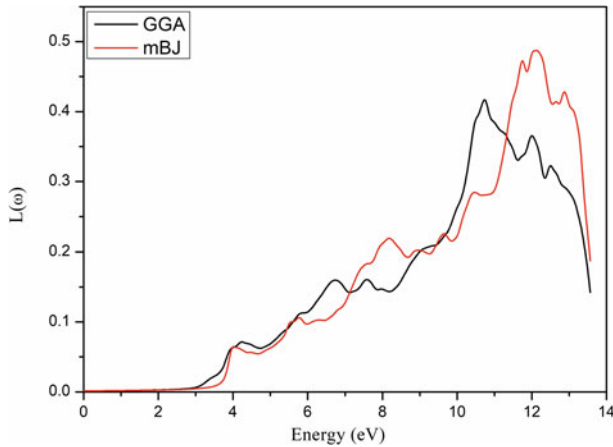


Fig. 10.8 The energy loss function for BaMgLaBiO₆ calculated with GGA and mBJ potentials

optoelectronic devices. The present results reveal the potential of double perovskite BaMgLaBiO₆ for lasers and power switching applications.

References

1. R.H. Mitchell, *Perovskites: Modern and Ancient*, vol. 7 (Almaz Press Thunder Bay, 2002)
2. A. Hossain, P. Bandyopadhyay, S. Roy, An overview of double perovskites A2B' B'' O6 with small ions at A site: Synthesis, structure and magnetic properties. *J. Alloy. Compd.* **740**, 414–427 (2018)
3. Y. Aharbil et al., Ferromagnetism and half metallicity induced by oxygen vacancies in the double perovskite BaSrNiWO6: DFT study. *Mater. Chem. Phys.* **183**, 588–594 (2016)
4. A.S. Thind et al., KBaTeBiO6: a lead-free, inorganic double-perovskite semiconductor for photovoltaic applications. *Chem. Mater.* **31**(13), 4769–4778 (2019)
5. D. Serrate, J. De Teresa, M. Ibarra, Double perovskites with ferromagnetism above room temperature. *J. Phys. Condens. Matter* **19**(2), 023201 (2006)
6. B. Fisher, et al., Variable range hopping in A₂MnReO₆ (A= Ca, Sr, Ba). *J. Appl. Phys.* **104**(3): 033716 (2008)
7. M. Sikora, et al., Evidence of unquenched Re orbital magnetic moment in AA' Fe Re O 6 double perovskites. *Appl. Phys. Lett.* **89**(6), 062509 (2006)
8. R. Cava, et al., Superconductivity near 30 K without copper: the Ba_{0.6}K_{0.4}BiO₃ perovskite. *Nature* **332**(6167), 814–816 (1988)
9. H. Das, M. De Raychaudhury, T. Saha-Dasgupta, Moderate to large magneto-optical signals in high T_C double perovskites. *Appl. Phys. Lett.* **92**(20), 201912 (2008)
10. M. Fiebig et al., Observation of coupled magnetic and electric domains. *Nature* **419**(6909), 818–820 (2002)
11. M.R. Filip et al., Band gaps of the lead-free halide double perovskites Cs₂BiAgCl₆ and Cs₂BiAgBr₆ from theory and experiment. *J. Phys. Chem. Lett.* **7**(13), 2579–2585 (2016)
12. A.H. Slavney et al., A bismuth-halide double perovskite with long carrier recombination lifetime for photovoltaic applications. *J. Am. Chem. Soc.* **138**(7), 2138–2141 (2016)
13. E.T. McClure et al., Cs₂AgBiX₆ (X= Br, Cl): new visible light absorbing, lead-free halide perovskite semiconductors. *Chem. Mater.* **28**(5), 1348–1354 (2016)

14. C. Marjerrison, et al., Magnetic ground states in the three O s 6+(5 d 2) double perovskites B a 2 M O s O 6 (M= Mg, Zn, and Cd) from Néel order to its suppression. *Phys. Rev. B* **94**(13), 134429 (2016)
15. S.A. Khandy, D.C. Gupta, Electronic structure, magnetism and thermoelectricity in layered perovskites: Sr₂SnMnO₆ and Sr₂SnFeO₆. *J. Magn. Magn. Mater.* **441**, 166–173 (2017)
16. C. Zhang et al., Three-dimensionally ordered macro-/mesoporous carbon loading sulfur as high-performance cathodes for lithium/sulfur batteries. *J. Alloy. Compd.* **714**, 126–132 (2017)
17. C. Thompson, et al., Frustrated magnetism in the double perovskite L a 2 LiOs O 6: A comparison with L a 2 LiRu O 6. *Phys. Rev. B* **93**(1), 014431 (2016)
18. Y.-P. Liu, H.-R. Fuh, Y.-K. Wang, Ab initio research on a new type of half-metallic double perovskites, A₂CrMO₆ (A= IVA group elements; M= Mo, Re and W). *Computation* **2**(1), 12–22 (2014)
19. G. Volonakis et al., Lead-free halide double perovskites via heterovalent substitution of noble metals. *J. Phys. Chem. Lett.* **7**(7), 1254–1259 (2016)
20. A. Kumar, M. Kumar, R. Singh, Magnetic, opto-electronic, and thermodynamic properties of half-metallic double perovskite oxide, Ba₂YbTaO₆: a density functional theory study. *J. Mater. Sci. Mater. Electron.* **32**(10), 12951–12965 (2021)
21. K.A. Parrey et al., Electronic structure, optical and transport properties of double perovskite La₂NbMnO₆: a theoretical understanding from DFT calculations. *J. Electron. Mater.* **47**(7), 3615–3621 (2018)
22. S. Halder et al., Electronic structure and electrical conduction by polaron hopping mechanism in A₂LuTaO₆ (A= Ba, Sr, Ca) double perovskite oxides. *Ceram. Int.* **43**(14), 11097–11108 (2017)
23. M. Cernea et al., Characterization of ferromagnetic double perovskite Sr₂FeMoO₆ prepared by various methods. *Ceram. Int.* **40**(8), 11601–11609 (2014)
24. N. Rammeh, Crystal structure, electronic and magnetic properties of double perovskite Ba₂FeWO₆: A combined experimental–theoretical study. *Physica B* **481**, 217–223 (2016)
25. P.N. Lekshmi et al., Structural, magnetic and dielectric properties of rare earth based double perovskites RE₂NiMnO₆ (RE= La, pr, Sm, Tb). *Physica B* **448**, 285–289 (2014)
26. P. Blaha, et al., wien2k. An augmented plane wave+ local orbitals program for calculating crystal properties (2001)
27. K. Schwarz, P. Blaha, G.K. Madsen, Electronic structure calculations of solids using the WIEN2k package for material sciences. *Comput. Phys. Commun.* **147**(1–2), 71–76 (2002)
28. J.P. Perdew, K. Burke, M. Ernzerhof, Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**(18), 3865 (1996)
29. J. Heyd, et al., Energy band gaps and lattice parameters evaluated with the Heyd-Scuseria-Ernzerhof screened hybrid functional. *J. Chem. Phys.* **123**(17), 174101 (2005)
30. F. Tran, P. Blaha, Accurate band gaps of semiconductors and insulators with a semilocal exchange–correlation potential. *Phys. Rev. Lett.* **102**(22), 226401 (2009)
31. M. Ameri, et al., First-principle investigations of structural, electronic and thermodynamic properties of CdS_{1-x}Se_x ternary alloys:(0.0 x 1.0). *Mater. Express* **4**(6), 521–532 (2014)
32. D.R. Penn, Wave-number-dependent dielectric function of semiconductors. *Phys. Rev.* **128**(5), 2093 (1962)
33. A.A. Khan, et al., DFT prediction of the structural, electronic, thermoelectric and optical properties of ternary pnictides MgBe₂X₂ (X= N, P, As, Sb, Bi): A novel analysis of beryllium with 2A- and 5B-Elements of the structure type CaAl₂Si₂. *Solid State Commun.* **300**, 113667 (2019)

Chapter 11

Conductivity of PANI/ZnO Nanocomposites



Siti Ashikin Hassan, Maytham Qabel Hamzah, and Mohd Arif Agam

Abstract The present work reports the synthetization and characterization of polymer Polyaniline/zinc oxide (PANI/ZnO) nanocomposites and was successfully fabricated by chemical oxidation method of ZnO via in situ polymerization method of PANI. Nowadays, the studies on conducting polymers are focus toward increasing the conductivity of the polymer that can be approached in many strategies such as doping the polymer with suitable nanoparticle in order to create highly packed and dense materials. Here, ZnO is doped into the polymer PANI with four different weight percentages of ZnO (2, 4, 6 and 8 wt %). The effect of the ZnO in terms of DC conductivity of the nanocomposites are observed through the four-point probe method. The highest conductivity reported is $1.25 \times 10^1 \text{ S cm}^{-1}$ at 8 weight percentage (wt %) of zinc oxide nanoparticles. The molecular structures had been studied using Fourier Transform Infrared Spectrometer (FTIR), and the spectra of (PANI/ZnO) nanocomposites were formed in emeraldine salt form which consists of the benzenoid and quinoid ring. The conductivity of the nanocomposites increases as the weight percentage of the nanocomposites increase. This high-conductive material can be used as a supercapacitor.

11.1 Introduction

The word polymer is a blend of two Greek words poly which means many and mers meaning units/parts. Thus, a polymer may be defined as a repeating small molecule (monomers) while in others may be branched or cross-linked. In other words, polymers are commonly applied in electrical and electronic applications such

S. A. Hassan · M. Q. Hamzah · M. A. Agam (✉)

Department of Physics, Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, 84600 Pagoh, Johor, Malaysia
e-mail: arif@uthm.edu.my

M. Q. Hamzah

The General Directorate of Education, Ministry of Education, Al-Muthanna Governorate, Republic of Iraq

as an insulator, where the advantages are taken due to their properties of very high resistivity. The polymers are broadly chosen because of their strength [1], flexibility [2], elasticity [3], stability, mouldability and ease of handling.

Polymers are insulators of electricity because they neither have a large number of charge carriers (free electrons or holes) nor an orbital system to make the charge carriers mobile (the conjugated backbone of the polymer), the two essential components for the polymer to manifest the charge-conduction process. That is why they have largely been used as a substitute for structural materials such as metals and alloys, wood, ceramics etc. Until the last few decades, polymers remained unsuccessful in replacing metals and semiconductors in electrical and electronic applications due to their insulating properties [4–6].

Many studies on conducting polymers are focus toward increasing the conductivity of the polymer that can be approached in many strategies such as synthesizing new molecular structure such as in Organic Semiconductor research (OSC) [7], doping the polymer with suitable nanoparticles [8] or restructuring of the polymer molecules to create highly packed and dense materials [9, 10]. Thus, to be classified as a conducting polymer, a polymer must possess the subsequent essential features:

1. Presence of extended conjugation which provides an excellent degree of delocalization of π -electrons within the molecules [11].
2. As pristine conjugated polymers do not contain characteristic charge transporters, charge transporters should tend an outward interaction, called doping.

Conducting polymer such as polyaniline (PANI) has special status among other conducting polymers due to its non-redox doping, good environmental stability and economic feasibility [12]. Besides, it become favoured conducting polymer because of its unique electrochemical properties, simple synthesis using low- cost chemicals and procedure, lightweight, high absorption coefficient in the visible light, and high mobility of charge carriers [13].

In past few years, organic and inorganic nanoparticle materials are let to mix into conducting polymer matrices in fabricating new polymer nanocomposite materials that have caught much attentions of researchers due to their wide range of application. This new polymer (organic or inorganic) (PO/PIO) composites are an interesting material as the properties of these composites can be tuned through the dopant concentration and polymer-organic/polymer-inorganic interactions. The second manipulation technique in tuning these PO/PIO properties, is by introducing the PO/PIO with manipulation techniques such as laser or radioactive radiation, chemical or thermal techniques [14].

Numbers of efforts have been made to successfully prepare the nanocomposites by chemical and electrochemical routes using nanostructured metal oxide namely copper oxide (CuO) [15], titanium dioxide (TiO₂) [16], silicon oxide (SiO₂) [16] and lead oxide (PbO) [17] because of their unique electrolytic, piezoelectric and phonic properties. Nanostructured zinc oxide has unique properties like high isoelectric point, an n-type semiconductor with direct band gap (3.37 eV) and high binding energy (60 meV), nontoxicity [18] and high electron transfer capability.

In this paper, PANI and its nanocomposites with zinc oxide nanoparticles (ZnO NPs) are fabricated by in situ oxidative polymerization of aniline monomer with ammonium peroxydisulphate (APS). The products are characterized by Fourier Transform Infrared (FTIR) spectra to identify chemical bond and four-point probe for electrical conductivity. The interaction between PANI and ZnO nanoparticles is investigated and the results show the doping effect of ZnO nanoparticles. The aim of the production of PANI/ZnO nanocomposite is in the future to be applied as a supercapacitor.

11.2 Methodology

Table 11.1 shows the list of chemicals used to synthesize the zinc oxide nanoparticles.

11.2.1 Preparation and Synthesis of Zinc Oxide Nanoparticles

ZnO NPs are synthesized by chemical oxidation method, zinc acetate dihydrate and ammonia solution are used as precursors and the soluble starch as stabilizing agent according to the procedure reported in the literature [19]. In this experiment, 2.5 g of soluble starch is dissolved in 500 ml of distilled water and the starch solution is kept under constant stirring and heating until becomes a clear solution.

Then, in another beaker, 11 g of zinc acetate dihydrate is dissolved in 100 ml of distilled water. The solution is stirred continuously for 10 min using a magnetic stirrer. Then, the solution is added into the starch solution and 2 ml of 1 M ammonia solution is added into the mixed solution drop by drop. The reaction of the solution is allowed to proceed under constant stirring for 2 h. After the complete of reaction, the solution is permitted to settle overnight and centrifuged at 5000 rpm for 10 min and the supernatant solution is discarded. Thus, the precipitate ZnO are washed three

Table 11.1 Chemicals used for zinc oxide nanoparticles

Chemicals	Molecular formula	Molar mass (g/mol)
Zinc acetate dihydrate	[Zn (CH ₃ CO ₂) ₂ .2H ₂ O]	219.50
Soluble starch	C ₁₂ H ₂₅ N O ₁₁	359.33
1 M Ammonia solution	NH ₄ -OH	35.04
Distilled water	H ₂ O	18.02
Deionized water	H ₂ O	18.02
Ethanol	C ₂ H ₆ O	46.07

times using distilled water and ethanol to remove the by-products and excessive starch which are bound with the nanoparticles. Afterwards, the precipitate are dried for overnight in an oven at 80 °C. A complete conversion of zinc hydroxide [Zn (OH)₂] into zinc oxide (ZnO) takes place during the drying process. The prepared ZnO NPs are used for further characterization and composites.

11.2.2 Preparation and Synthesis of PANI/ZnO Nanocomposite

Table 11.2 shows the chemicals used to synthesize the nanocomposites.

PANI/ZnO NCs are prepared by in situ polymerization of aniline with four different weight percentages of ZnO: 2, 4, 6 and 8 wt % to produce NC1, NC2, NC3, and NC4, respectively.

0.042 g (2 wt%) of the above synthesized ZnO NPs are poured into a beaker containing a solution of 100 ml of 1 M HCl and 2 ml of aniline monomer, followed by the sonication of the solution for 1 h at 40 °C. The sonication is to ensure the ZnO NPs and the aniline are dissolved homogeneously in the solution. After that, the solution is stirring constantly in an ice bath using the magnetic stirrer for 1 h at 5 °C. In another beaker, 4.9 g of APS is poured into 50 ml of pre-cooled 1 M HCl. Then, the reaction is continued by adding dropwise the APS and HCl solution. The stirring is continued at 20 °C for 5 h long until the polymerization of aniline and ZnO NPs homogeneously stirred. The process proceeds with the centrifugation of emerald black solution with deionized water, methanol and acetone at 10 000 rpm for 10 min to remove unreacted excess products. The precipitate obtained is dried in an oven overnight at 80 °C and grinded with pastel and mortar. The steps followed by another weight percentage (4, 6 and 8 wt %).

Table 11.2 Chemicals used for nanocomposites

Chemicals	Molecular formula	Molar mass (g/mol)
Aniline	C ₆ H ₅ N H ₂	93.13
Ammonium perodisulfate	(NH ₄) ₂ S ₂ O ₈	228.201
Hydrochloric acid	HCl	36.458
Zinc acetate dihydrate	[Zn (CH ₃ CO ₂) ₂ .2H ₂ O]	219.50
Soluble starch	C ₁₂ H ₂₅ N O ₁₁	359.33
1 M Ammonia solution	NH ₄ -OH	35.04

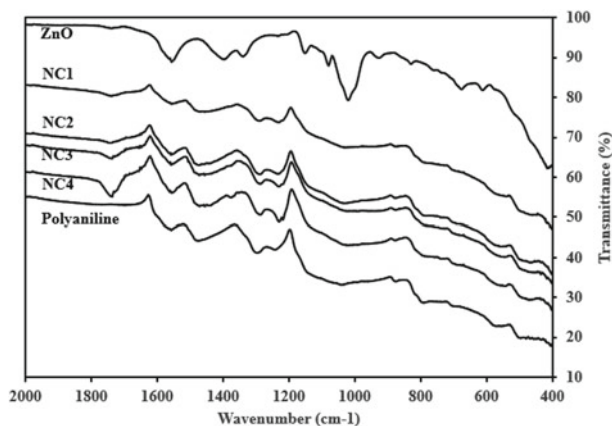


Fig. 11.1 FTIR spectra of PANI, ZnO, and NCs (NC1, NC2, NC3, and NC4)

11.3 Results and Discussion

11.3.1 FTIR Spectroscopy

FTIR spectrum is the feature of a particular compound that gives the information about its functional groups, molecular geometry and inter/intramolecular interactions. FTIR spectra of the prepared ZnO, polyaniline and the nanocomposites are shown in Fig. 11.1. The spectrum of ZnO shows peaks at 1555 cm^{-1} which are assigned to adsorbed water at the ZnO surface [20]. In the polyaniline spectrum, the bands appeared at 521 cm^{-1} and 831 cm^{-1} are due to C–N–C of the aromatic ring and C–H in the plane of the benzenoid ring. A characteristic absorption peak of polyaniline that emerged at 1511 cm^{-1} is assigned to the C=C bond in the benzenoid ring, which gives evidence of the presence of doped PANI structures. The presence of the benzenoid and quinoid units is evidence of the emeraldine form of PANI. The FTIR spectra of the polyaniline/zinc oxide nanocomposites show the same characteristics peaks of the PANI.

11.3.2 Electrical Conductivity

The variation of DC conductivity of the prepared nanocomposites with a different weight percentage of zinc oxide is shown in the Table 11.3. The figure shows that the electrical transport (conduction) is represented by the convergence of ZnO in the nanocomposite as the ZnO acts a significant part in the conduction system. The DC electrical conductivity of the nanocomposites room temperature, is measured using a four-point probe and listed in Table 11.1. From this table it is clear that the

Table 11.3 DC conductivity (σ_{dc}) of PANI, ZnO NPs, and PANI/ZnO NCs at room temperature

Nanocomposite weight percentage (%)	0 (pure pani)	2	4	6	8
Conductivity (S/cm)	5.18	6.20	7.45	10.40	12.50

conductivity of pure polyaniline is 5.18 S/cm and increases with increasing weight percentage of ZnO NPs to (6.20, 7.45, 1.04×10^1 , 1.25×10^1) at (2 wt %, 4 wt %, 6 wt % 8 wt %) respectively. The conductivity of the PANI/ZnO nanocomposite increase with increasing the weight percentage of the nanocomposite from 2 wt% to 8 wt%. The conductivity of the conjugated polymer fundamentally relies upon its conjugation length [21].

The pure PANI conductivity is related to the electron transportation conductivity mechanism of its conjugated system. The increase in the electrical conductivity with the further increase in the ZnO content in each nanocomposite could be attributed to the charge carries of the ZnO nanoparticles, which is based on electrons and holes and also the ability of the nanoparticles to fill in the tiny blanks in the polymer matrix [22].

11.4 Conclusion

For FTIR, the nanocomposite with an increasing weight percentage of ZnO, some vibrational bands are shifted to higher wavenumbers. Due to the H-bonding interaction between polyaniline chains and hydrochloric acid, and the hydrogen bond between the amine and imine nitrogen site in the PANI chain. The FTIR data revealed that the spectra of the nanocomposites are formed in the emeraldine salt form which consists of a benzenoid and quinoid form. An increase in the electrical conductivity indicates that there is a transition.

Acknowledgements This work was funded by the Fundamental Research Grant Scheme awarded by the Ministry of Higher Education Malaysia [FRGS/1/2019/STG07/UTHM/02/5 (FRGS K171)] and TIER 1 (H218) and GPPS (H423), awarded by University Tun Husain Onn Malaysia (UTHM), Malaysia.

References

1. S.K.S. Basha, G.S. Sundari, K.V. Kumar, M.C. Rao, Preparation and dielectric properties of PVP-based polymer electrolyte films for solid-state battery application. *Polym. Bull.* **75**(3), 925–945 (2018). <https://doi.org/10.1007/s00289-017-2072-5>
2. J. Li et al., Flexible, conductive, porous, fibrillar polymer-gold nanocomposites with enhanced electromagnetic interference shielding and mechanical properties. *J. Mater. Chem. C* **5**(5), 1095–1105 (2017). <https://doi.org/10.1039/c6tc04780g>

3. A.I. Isayev, Generalized characterization of relaxation properties and high elasticity of polymeric systems. *J. Polym. Sci. Part A-2 Polym. Phys.* **11**(11), 2123–2133 (1973). <https://doi.org/10.1002/pol.1973.180111104>
4. C. Laurent, Optical prebreakdown warnings in insulating polymers. *IEEE Electr. Insul. Mag.* **15**(2), 5–13 (1999). <https://doi.org/10.1109/57.753926>
5. D.L. Zhang et al., High thermal conductivity and excellent electrical insulation performance in double-percolated three-phase polymer nanocomposites. *Compos. Sci. Technol.* **144**, 36–42 (2017). <https://doi.org/10.1016/j.compscitech.2017.02.022>
6. G. Lu et al., Enhanced charge transportation in semiconducting polymer/insulating polymer composites: The role of an interpenetrating Bulk interface. *Adv. Funct. Mater.* **20**(11), 1714–1720 (2010). <https://doi.org/10.1002/adfm.200902281>
7. Z. Luo et al., Side-chain effects on energy-level modulation and device performance of organic semiconductor acceptors in organic solar cells. *ACS Appl. Mater. Interfaces* **9**(39), 34146–34152 (2017). <https://doi.org/10.1021/acsami.7b10275>
8. S. Ibrahim, U. Muhammad, A.A. Muhammad, Humidity Sensing Study of Polyaniline/Copper Oxide Nanocomposites. *Int. J. Adv. Acad. Res.* **4**(5), 49–61 (2018)
9. S.M. Swick et al., Closely packed, low reorganization energy π -extended postfullerene acceptors for efficient polymer solar cells. *Proc. Natl. Acad. Sci. U. S. A.* **115**(36), E8341–E8348 (2018). <https://doi.org/10.1073/pnas.1807535115>
10. T.H. Le, Y. Kim, H. Yoon, Electrical and electrochemical properties of conducting polymers. *Polymers (Basel)* **9**(4) (2017). <https://doi.org/10.3390/polym9040150>
11. A.K. Mishra, Conducting polymers: concepts and applications. *J. At. Mol. Condens. Nano Phys.* **5**(2), 159–193 (2018). <https://doi.org/10.26713/jamcnp.v5i2.842>
12. D. Ghosh, S. Giri, S. Kalra, C.K. Das, Synthesis and characterisations of TiO₂ coated multiwalled carbon nanotubes/graphene/polyaniline nanocomposite for supercapacitor applications. *Open J. Appl. Sci.* **02**(02), 70–77 (2012). <https://doi.org/10.4236/ojapps.2012.22009>
13. T.K. Das, S. Prusty, Review on conducting polymers and their applications. *Polym. Plast. Technol. Eng.* **51**(14), 1487–1500 (2012). <https://doi.org/10.1080/03602559.2012.710697>
14. Fitria, 濟無No Title No Title. *J. Chem. Inf. Model.* **53**(9), 1689–1699 (2013)
15. S.P. Ashokkumar, H. Vijeth, L. Yesappa, M. Niranjana, M. Vandana, H. Devendrappa, Electrochemically synthesized polyaniline/copper oxide nano composites: To study optical band gap and electrochemical performance for energy storage devices. *Inorg. Chem. Commun.* **115**(March), 107865 (2020). <https://doi.org/10.1016/j.inoche.2020.107865>
16. S.O. Mezan, et al., Synthesis, characterization, and properties of polystyrene/SiO₂ nanocomposite via sol-gel process. *AIP Conf. Proc.*, **2151**(August) (2019). <https://doi.org/10.1063/1.5124664>
17. J.H. Syed Abusalehamadnabirqudri, M. Mir Mohiuddin, M. Ahmed, Thermal and Sensing Properties Polyaniline/Lead Oxide Nanocomposites **8**(May), 45–47 (2018). <https://doi.org/10.9790/9622-0805044547>
18. A. Kolodziejczak-Radzimska, T. Jesionowski, Zinc oxide-from synthesis to application: A review. *Materials (Basel)* **7**(4), 2833–2881 (2014). <https://doi.org/10.3390/ma7042833>
19. S.M. Giripunje, J. Ghushie, Nanocomposite of polyaniline and ZnO: Preparation, characterization, optical and electrical properties. *J. Nano Res.* **24**, 123–132 (2013). <https://doi.org/10.4028/www.scientific.net/JNanoR.24.123>
20. A. Mostafaei, A. Zolriasatein, Synthesis and characterization of conducting polyaniline nanocomposites containing ZnO nanorods. *Prog. Nat. Sci. Mater. Int.* **22**(4), 273–280 (2012). <https://doi.org/10.1016/j.pnsc.2012.07.002>
21. R.M. Mohsen, S.M.M. Morsi, M.M. Selim, A.M. Ghoneim, H.M. El-Sherif, Electrical, thermal, morphological, and antibacterial studies of synthesized polyaniline/zinc oxide nanocomposites. *Polym. Bull.* **76**(1), 1–21 (2019). <https://doi.org/10.1007/s00289-018-2348-4>
22. Y.M. Ahmed, S.M. Hassan, *Effect of ZnO Particle Size on A. C. Electrical Properties of Prepared Polyaniline/ZnO Composites*, no. May, pp. 34–38 (2020)

Chapter 12

A Mach–Zehnder Ring Resonator Layout for Label Free Detection of Diabetes



Parisa Sanati, Mahdi Bahadoran, and Mohd Kamarulzaki Mustafa

Abstract Most diabetic patients argue about the high price of the disposable test card together with an unpleasant and painful fasting blood test that must be repeated daily. In this work, a Mach–Zehnder interferometer was combined with a small size ring resonator system, including two half-circular racetrack ring resonators to set up a Mach–Zehnder ring resonator (MZRR) system for low cost and painless process of diabetes detection. In this work, patient's tear was exploited instead of using blood samples to get rid of daily painful blood taking test. The optical transfer function of MZRR is derived using the signal flow graph method. The MZRR works based on a variation of the refractive indices of the test samples and any change in the refractive index of the top cladding sensing layer can cause a specific shift in output light resonance peaks. The sensitivity of the MZRR sensor for the detection of diabetes is measured to be 360 nm/RIU, which leads to a detection limit of 5.5×10^{-5} RIU. The advantages of the MZRR sensor are in high accuracy, the quick approach of detection and a painless process, which make the MZRR sensor a proper lab-on-chip device for health monitoring of diabetic patients.

12.1 Introduction

The design of the label-free biosensors has been a subject of interest in medical diagnostics, pharmaceutical industries and monitoring of the environment. The design of miniaturized and high sensitive sensors has been receiving great attention in the

P. Sanati

Burn and Wound Healing Research Center, Shiraz University of Medical Sciences, 71345-1978 Shiraz, Iran

M. Bahadoran

Department of Physics, Shiraz University of Technology, 31371555 Shiraz, Fars, Iran

M. K. Mustafa (✉)

Department of Physics and Chemistry, Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, Pagoh Educational Hub, 86400 Muar, Johor, Malaysia
e-mail: kamarulz@uthm.edu.my

development of integrated optical biosensors [1]. Amongst the various types of optical sensors, silicon-based sensors have received a wide range of interests in the integrated electro-optic microfluidic and Lab-on-chip tools, thanks to their unique features like compatibility with the standard Complementary Metal-Oxide Semiconductor, (CMOS) fabrication procedure, no sensitivity to an electromagnetic wave, possibility to design a fabricate in micro size, low cost and high mode confinement [2–4]. These characteristics of silicon-based sensors deliver clinical diagnostics and fully-automated biological analysis that is desirable in the pharmaceutical, medical and food industries [5]. Generally, the two approaches of invasive and non-invasive have been used for the identification of biological cells. The label-based method (fluorescence-based approach) is an invasive method. The no marker method is known as non-invasive as no labels are required to be attached to the bio-cells, thus, the cell's information remains intact [6]. Label-free detection (LFD) delivers straightforward, natural and highly sensitive detection of bio-cells free from adding markers or impurities to the target sample. The light interaction with the test sample (bio-cells) that cause a change in the output light response is the key parameter in Label-free sensors [7]. Nowadays, label-free detection has been applied in interferometric sensors [8, 9], photonic crystal sensors, chemical sensors [10], electrochemical biosensor [11] and refractive index-based sensors [12, 13]. In the label-free refractive index sensors [14, 15], any change in the refractive index of the test sample contributes a change in the output light response. Among the conventional label-free refractive index sensors, ring resonator-based devices [16–20] have received a considerable attention in sensing applications due to their specific features like high Q-factor, small size, low power consumption, low cost and easy fabrication process and high efficiency [21–23]. In this work, we applied a layout of the ring resonator for quick detection of diabetes.

Diabetes step from an excess of sugar in the blood. A high level of sugar in the blood can cause major health concerns^{24, 25}. Glucose is known as the brain's primary fuel source and is a main source of energy for cells that makes up the muscles and tissues and is essential to human health. The underlying cause of diabetes differs depending on the kind. The type 1, type 2, and gestational diabetes are the three basic kinds of diabetes²⁶. Type 1 diabetes is an autoimmune disease in which the pancreas fails to make insulin. It can strike individuals at any age. Diabetes, type 1 is the most common kind of diabetes. The failure of the body to produce insulin, the hormone that unlocks the body's cells, allowing glucose to enter and fuel them, causes diabetes. Insulin resistance is the cause of type 2 diabetes. Type 2 diabetes known as a frequent illness in which the blood sugar (glucose) level rises to highly dangerous levels. During pregnancy, the placenta produces hormones that increase blood glucose levels. The pancreas can usually produce enough insulin to manage the situation. However, if the body does not produce enough insulin or does not use it properly, blood sugar levels rise, resulting in gestational diabetes²⁷. The diabetic patients suffer from frequent urination, feeling very thirsty, exhaustion and may feel significant heart, eye and nerve problems. Biochemical methods are the most prevalent way to measure blood glucose. In terms of the assay concept, calibration procedure, and result expression, these approaches are complicated, inconvenient and

expensive. There exist some electro-chemical tools in the markets, which work based on a sample of finger blood and disposable measuring stick. Sometimes diabetic patients argue about the uncomfortable blood test by applying a lancet into their finger [28] and some patients in poor countries demand for low-cost approaches [24]. Due to the inadequacies of existing approaches, a novel method with high accuracy, fast detection, low cost and low pain is required. Here, we took advantage of the small size Mach–Zehnder resonator system, including two half-circle ring resonator and two straight bus waveguide sections that made a asymmetric ring resonator system, which fulfils the expectations of low price, fast process of detection, and high precision. Besides, we change the test sample from blood to tear of diabetic patients to get rid of daily painful blood taking test.

In this work, a combined Mach–Zehnder interferometer with half-ring resonator system is selected in a set up a Mach–Zehnder ring resonator (MZRR) system for label-free detection of diabetes. The MZRR works based on a variation of the refractive indices of the test samples. Theoretically, any change in the index of refraction of the superstrate layer of the sensing waveguide can cause a definite shift in output resonance peaks. The merits of our proposed sensor are in low cost of fabrication, high sensitivity and quick approach of detection without any pain, which make the MZRR sensor a proper lab-on-chip device for health monitoring of diabetic patients.

12.2 Methodology

The free spectral range (*FSR*) is defined as the wavelength difference between two sequential resonance peaks that theoretically can be calculated by $FSR = \lambda_0^2/n_g L$ [15, 25]. Here λ_0 shows the wavelength between sequential resonance peaks, n_g is the group index of optical medium and L represents the optical length. Applying a narrow line width tunable laser together with using high precision optical spectrum analyzer (OSA) [26] are conventional so as to improve of sensitivity. The ratio of output to input light power can be determined using the signal flow graph (SFG) method [25, 27]. The SFG method makes use of Mason's rule [28, 29] to calculate optical transfer function (OTF) of photonic tools. Based on the Mason's rule, the OTF of an optical device with E_{in} and E_{out} light fields at the input and output ports, respectively is given by [27, 30], The volume editors, usually the program chairs, will be your main points of contact for the preparation of the volume.

$$OTF = \frac{E_{out}}{E_{in}} = \frac{\sum_{m=1}^N T_m \Delta_m}{1 - \sum_m L_m + \sum_{m \neq g} L_m L_g - \dots} \quad (12.1)$$

where T_m is the m th forward trajectory from E_{in} to E_{out} ports, Δ_m is defined as the delta determinant for m th track [31, 32] and N denotes for the total number of one-way paths from E_{in} to E_{out} ports. Here, L_i represents the closed loop existed in photonics nodes [23, 33] (red dots in Fig. 12.1b). Configuration of Mach–Zehnder

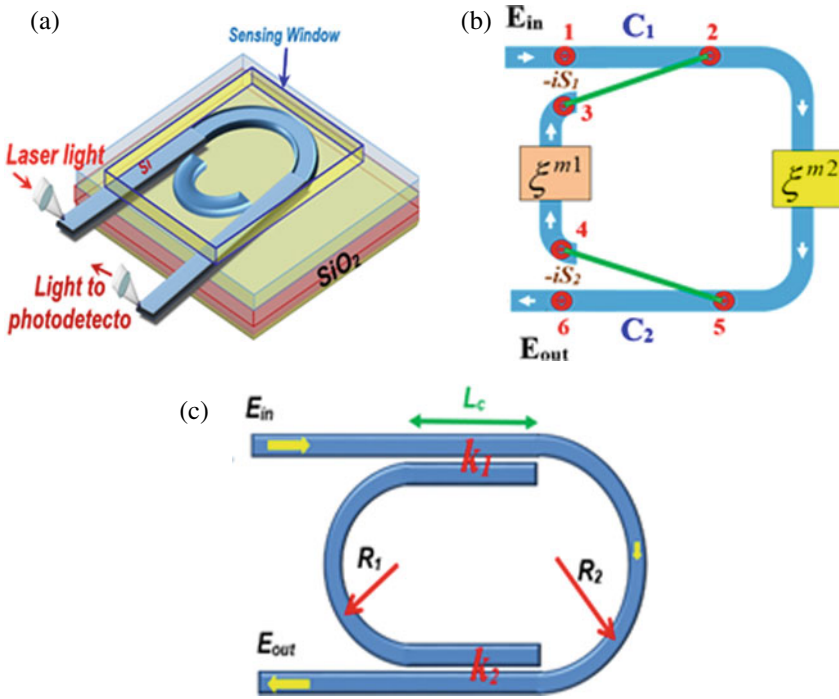


Fig. 12.1 A generic layout of MZRR sensor **a** the sensor configuration **b** SFG diagram **c** the waveguide arrangement

ring resonator (MZRR) is displayed in Fig. 12.1a and the SFG diagram for MZRR is given in Fig. 12.1b. In each coupler, $S_n = \sqrt{(1 - \gamma_n)k_n}$ shows the fraction of light pulse evanescent over the waveguides [33–35] and $C_n = \sqrt{(1 - \gamma_n)(1 - k_n)}$ denotes to the portion of input light to the coupler that travels via one waveguide [14]. Here, k_n shows the coupling coefficient and γ_n shows the coupling intensity loss for each coupler. The loss-unit delay parameter in SFG diagram is shown by $\varepsilon^m = \exp(-m\alpha L/2) \exp(-imk_n n_g L)$ [36–39] that m is the resonance mode number, α shows the internal loss factor of optical path, L is the circumference of the close loop, k represents the wave number and n_g shows the group refractive index of ring waveguide.

According to the Mason rule, all of the closed loops and forward transition paths from the E_{in} to E_{out} ports must be identified, then applied in Eq. (12.1). For MZRR layout, there exists only one closed loop passes through the photonics nodes of $2 \rightarrow 5 \rightarrow 4 \rightarrow 3 \rightarrow 2$, so the transmittance function of this loop is $L(MZRR) = S_1 S_2 \xi^{m_1+m_2}$. Since half part of close loop includes m_2 and the other part includes m_1 , ($m_1 + m_2 = 1$), thus $L(MZRR) = S_1 S_2 \xi$. The input light can solely transmit via photonics nodes of $1 \rightarrow 2 \rightarrow 5 \rightarrow 6$ to reach the output port. Thus, the transmit function of this unique forward path is $T_1(MZRR) = C_1 C_2 \xi^{m_2}$. As this track touched the $L(MZRR)$ loop, the delta determinant of this transmitting

track is $\Delta_1(MZRR) = 1$. Applying T_1 , Δ_1 and L_1 into the Eq. (12.1), the optical transfer function for Mach–Zehnder ring resonator can be calculated as in Eq. 12.2:

$$OTF_{MZRR} = \frac{E_{out}}{E_{in}} = \frac{C_1 C_2 \xi^{m^2}}{1 - S_1 S_2 \xi} \quad (12.2)$$

The normalized intensity of light via the MZRR system can be calculated by a dot product of optical transfer function to its complex conjugate as $OTF \cdot OTF^*$.

12.3 Simulated Result and Discussion

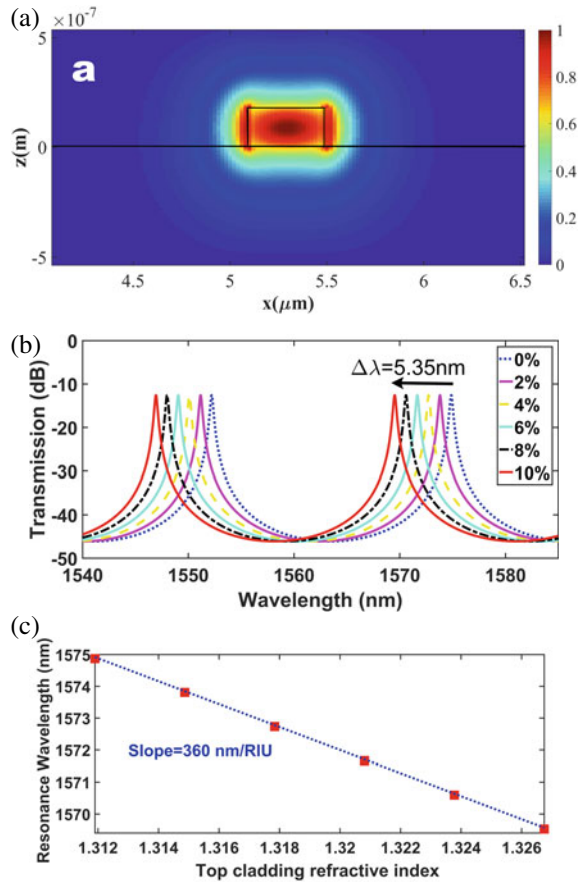
As demonstrated in Fig. 12.1a, a silicon-on-insulator (SOI) waveguide is used in the MZRR system. The structure included a 2 μm height of SiO₂ substrate, a silicon core with the height and width of 180 nm and 400 nm, respectively that were placed on silica substrate. As shown in Fig. 12.1c, the asymmetric MZRR has two straight sections, Lc, as large as 0.5 μm and two half circles. The larger half circle has the radius of 4.2 μm and the radius of smaller one is 3.7 μm . The coupling gaps are adjusted on 100 nm. A tunable laser with a center wavelength of 1550 nm was applied as input light sources and optical spectrum analyzer (OSA) was exploited as an output light detector. Applying various solutions on the top layer of the silicon waveguide can change the group index of the waveguide and it causes a wavelength shift in the output resonance peaks. In the first step, the MZRR sensor is calibrated by applying different concentrations of sugar solution. The reflective index of various concentrations (mass %) of sugar solution, C, in the wavelength of 1550 nm is given by Eq. (12.3) [32, 40],

$$n = 1.3119 + 0.14849 \times C\% \quad (12.3)$$

Based on Eq. (12.3), the refractive indices of 1.31190, 1.31487, 1.31784, 1.32081, 1.32378 and 1.32675 are correspond to the sugar solutions with the concentration of 0%, 2%, 4%, 6%, 8% and 10%, respectively. Implementing these solutions (0 to 10%) on the window sensor leads to the group indices of 4.282559, 4.279663, 4.276748, 4.273834, 4.270922 and 4.268013 respectively.

In the microring-based biosensor, the propagating light through the waveguide core is required to evanescently couple to the top cladding part of the waveguide and interact with the molecules of the test sample [41]. The interaction of the evanescent light with the bio-cells contributes to a change in the effective index and group index of the sensing waveguide and consequently can make a spectral shift in the output light [12, 13]. The transfer electric mode (the lowest order mode) of light through the ring waveguide is displayed in Fig. 12.2a. The core of the waveguide has the maximum light intensity (dark red) and the evanescent light is shown by the yellow and blue halos around the core. As illustrated in Fig. 12.2a, the evanescent light from the core would penetrate almost 500 nm into the superstrate part of waveguide that

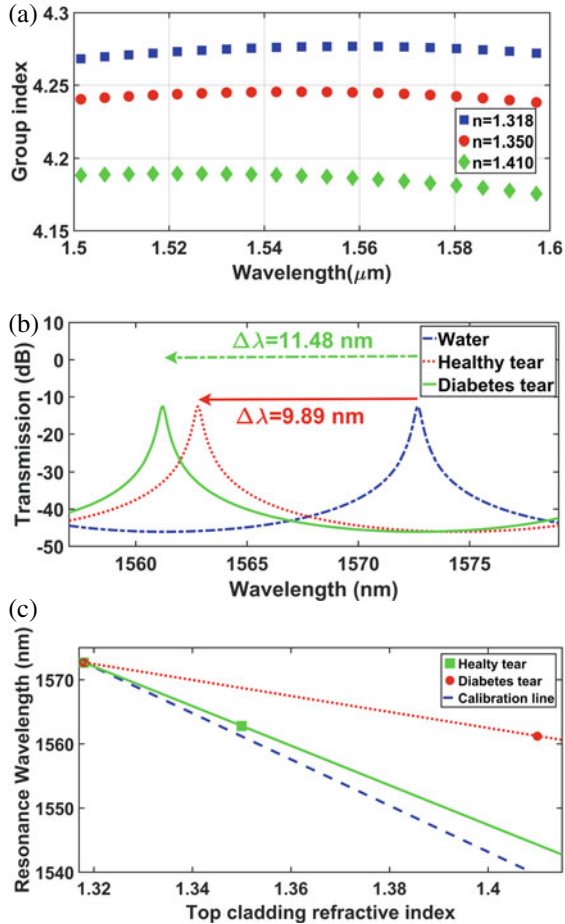
Fig. 12.2 **a** Transfer magnetic mode of light via sensor waveguide. **b** Optical transmission versus wavelength for the MZRR sensor from different concentrations of sugar solution. The optical parameters used in simulation of sensor are as follows: $k_1 = k_2 = 0.9$; $\gamma_1 = \gamma_2 = 0.01$; $\alpha = 5$ dB/cm **(c)** calibrated sensing line for MZRR



is almost three times larger than the core height. It means that this sensor can be used for the biological samples having submicron size of components. Applying different concentrations of water-alcohols [42, 43] water-salt [38, 44–46] or water-sugar [39, 47] have been considered as the benchmark for the calibration of the microring-based sensors.

Applying distilled water on the MZRR sensing window provides the free spectral range of 23 nm with a full width at half maximum as narrow as 330 pm that corresponds to the high Q factor of 4772. The wavelength response of the MZRR sensor for applied 0% to 10% concentrations of sugar solution on the sensing window is simulated in Fig. 12.2b. This change in the sample concentration brings about a group index variation that leads to a wavelength shift in the resonant peaks. For a change of 0.014 in the top cladding refractive index change, the resonant peaks in the wavelengths about 1570 nm have undergone the wavelength shifts of 5.35 nm that led to a sensitivity of 360 nm/RIU. The sensitivity of the refractive index-based sensor are measured by $\Delta\lambda_{peak}/\Delta n_{test}$ [13, 40, 48] that $\Delta\lambda_{peak}$ denotes a shift in the

Fig. 12.3 **a** the variation of group index of the sensing waveguide versus wavelength for applying water (blue square), healthy tear (red circles) and diabetes tears (black diamonds) **b** light transmission via MZRR versus wavelength for diabetes detection **c** sensing line for detection of diabetes



resonance peaks and Δn_{test} is the change in the refractive index of the test sample. Compare to the sensitivities reported by a single microring resonator-based sensors given in Table 12.1, current MZRR sensor has smaller size, higher sensitivity and high Q factor that is due to its mixed-design with Mach–Zehnder interferometer.

For detection of diabetes, three samples, including water (base line), a healthy tear sample and a tear sample of diabetes were applied on the sensing window of MZRR sensor. The refractive indices of water, healthy tear and diabetic patient’s tear are 1.318 [54, 55], 1.350 and 1.410, respectively. The refractive index of bio-cells can be measured using spectral attenuation data. The change in the group index of the sensing waveguide against the wavelength spectrum of 1500 nm-1600 nm for three solution samples of water, healthy and diabetes tears is shown in Fig. 12.3a. The group refractive index of these three samples is 4.277, 4.245 and 4.188 respectively in the wavelength of 1550 nm. The transmission light shows the blue-shift for tear samples rather than water baseline. The wavelength shifts as high as 9.89 nm and

Table 12.1 Comparison of different single ring resonator-based sensors

Sensor configuration	Sensitivity nm/RIU	Highlight	Size of sensor μm^2	Ref
Slot-waveguide-based SOI ring resonator	298	$DL = 4.2 \times 10^{-5}$ RIU	13×10	[49]
SOI racetrack ring resonator	70	10^{-5} RIU	10×10	[50]
SOI racetrack ring resonator	70	$Q = 2340$	12×8	[51]
Hybrid plasmonic waveguide micro-ring resonator	180	2.917×10^{-4}	2×2	[52]
SOI-photonic crystal ring resonator (PCRR)	111	$Q = 3700$	–	[53]
Mach-Zehnder Ring Resonator	360	5.5×10^{-5} RIU $Q = 4771$	8.5×8.5	Present work

11.48 nm were realized for diabetes and healthy tear samples, respectively as shown in Fig. 12.3b. The slope of each sensing line can give the sensitivity. Simulated results show that the tear healthy sample brings about a line almost near to the calibration line (blue-dash line). It means that the healthy samples will be located on the green line. The samples having diabetic tear generated a red dotted line with a slope of 124.65 nm/RIU. This difference in slope is due to the difference between the refractive indices of sugar solutions (used in the calibration process) and the tear samples that brings about a modification in the group indices of each sample. Indeed, achieving a difference between the healthy and infected samples is commonly expected in ring-based sensors [42, 43]. One can easily check their diabetic status by adding a drop of tear into the sensor and considering the locus of the test sample in the green and red lines in Fig. 12.3c. The tear from healthy person will be realized through the green line while the diabetes samples will be appeared along the red line. Totally, the MZRR sensor can ease and accelerate the diagnostic of diabetes patients. This sensor potentially can be used for as a disposable lab-on-chip device for daily monitor of blood sugar.

12.4 Conclusion

A novel approach is successfully employed for a low cost and painless detection of diabetic patients using patient tear instead of patient blood conventionally. Refractive index-based sensor designed using a Mach–Zehnder interferometer formed by half ring resonators for label free detection of diabetes was developed and able to differentiate between healthy and diabetic tear. The patient’s tear is used as a test

sample instead of using diabetic fingertip blood which is not painful and simulated using Matlab software. The sensor has the sensitivity and detection limit of 360 nm/RIU and 5.5×10^{-5} RIU, respectively, that can be potentially applied as lab-on-chip devices for daily monitoring of diabetic patients. Overall, the MZRR proposed sensor gives several advantages, including a quick approach of detection, small size, high sensitivity and low-cost of fabrication that can be actively used in health monitoring of diabetes.

Acknowledgements The authors would like to acknowledge for the research facilities of Shiraz University of Technology, Iran and Photonics Device and Sensor Research Centre (PDSR), Department of Physics and Chemistry, Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia.

References

1. P. Azuelos, P. Girault, N. Lorrain, L. Poffo, M. Guendouz, M. Thual, J. Lemaitre, P. Pirasteh, I. Hardy, J. Charrier, *Opt. Express* **25**(24), 30799–30806 (2017)
2. R. Guider, D. Gandolfi, T. Chalyan, L. Pasquardini, A. Samusenko, C. Pederzoli, G. Pucker, L. Pavesi, *Sens. Bio-Sens. Res.* **6**, 99–102 (2015)
3. B. de Oliveira, W. Daniel, A. Herbster, Presented at the 2019 SBFoton International Optics and Photonics Conference (SBFoton IOPC) (2019) (unpublished).
4. M.-M. Babakhani-Fard, A. Daraei, N. Hatefi-Kargan, *Eur. Phys. J. Plus* **135**(3), 1–13 (2020)
5. S.T. Fard, S.M. Grist, V. Donzella, S.A. Schmidt, J. Flueckiger, X. Wang, W. Shi, A. Millspaugh, M. Webb, D. M. Ratner, Presented at the Silicon Photonics VIII (2013) (unpublished)
6. H. Gu, H. Gong, C. Wang, X. Sun, X. Wang, Y. Yi, C. Chen, F. Wang, D. Zhang, *Sensors* **19**(22), 5038 (2019)
7. S. Sahu, J. Ali, G. Singh, *Optica Applicata* **48**(1) (2018)
8. J. Gao, T. Gao, Y.Y. Li, M.J. Sailor, *Langmuir* **18**(6), 2229–2233 (2002)
9. K. Schmitt, B. Schirmer, C. Hoffmann, A. Brandenburg, P. Meyrueis, *Biosens. Bioelectron.* **22**(11), 2591–2597 (2007)
10. H. Zhu, J.D. Suter, X. Fan, in *Optical guided-wave chemical and biosensors II* (Springer, 2010), pp. 259–279
11. X. Liu, R. Zhang, C. Rejeeth, S. Rana, D.D. Gurav, K. Qian, Presented at the 2018 IEEE International conference on manipulation, manufacturing and measurement on the nanoscale (3M-NANO) (2018) (unpublished)
12. P. Sanati, S.-S. Hashemi, M. Bahadoran, A.A. Babadi, E. Akbari, *Silicon* (2021)
13. P. Sanati, A. Shafiee, M. Bahadoran, E. Akbari, *Eur. Phys. J. Plus* **135**(11), 869 (2020)
14. A. Mohamad, M. Bahadoran, S. Daud, K.T. Chaudhary, M. Aziz, M. Jalil, J. Ali, P. Yupapin, *Jurnal Teknologi* **78**(3), 61–65 (2016)
15. A. Shafiee, M. Bahadoran, P. Yupapin, *Appl. Opt.* **58**(30), 8167–8173 (2019)
16. A.K. Seyfari, M. Bahadoran, A. Aghili, *OQE* **52**(9), 408 (2020)
17. M. Aziz, M. Affandi, S. Daud, M. Bahadoran, K. Chaudhary, J. Ali, *Jurnal Teknologi* **78**(3), 293–300 (2016)
18. I. Amiri, A. Afroozeh, M. Bahadoran, *Chin. Phys. Lett.* **28**(10), 104205 (2011)
19. M. Bahadoran, M.A. Jalil, J. Ali, I.S. Amiri, P. Yupapin, *OQE* **52**(216), 1–14 (2020)
20. I. Amiri, R. Zakaria, T. Anwar, M. Bahadoran, D. Vigneswaran, P. Yupapin, *Results Phys.* **11**, 1087–1093 (2018)
21. T. Claes, W. Bogaerts, P. Bienstman, *Opt. Express* **18**(22), 22747–22761 (2010)

22. M. Bahadoran, A.F.A. Noorden, K. Chaudhary, F.S. Mohajer, M.S. Aziz, S. Hashim, J. Ali, P. Yupapin, *Sensors* **14**(7), 12885–12899 (2014)
23. C. Sirawattananon, M. Bahadoran, J. Ali, S. Mitatha, P.P. Yupapin, *Nanotechnol. IEEE Trans.* **11**(4), 707–712 (2012)
24. A. Misra, H. Gopalan, R. Jayawardena, A.P. Hills, M. Soares, A.A. Reza-Albarrán, K.L. Ramaiya, *J. Diabetes* **11**(7), 522–539 (2019)
25. A.R. Lubis, S. Prayudani, M. Lubis, Presented at the Journal of Physics: Conference Series (2019) (unpublished)
26. D. McCay, S.C. Banerjee, L.M. Ghosal, M.M. Dutta, C. Ray, *Indian J. Med. Res.* **125**(3), C508 (2007)
27. J. Yamamoto, L. Donovan, K. Mohammad, S. Wood, *Diabet. Med.* **37**(1), 138–146 (2020)
28. J.A. Rowan, W.M. Hague, W. Gao, M.R. Battin, M.P. Moore, *N. Engl. J. Med.* **358**(19), 2003–2015 (2008)
29. C. Fernandez, *Labiotech* **23** (2018)
30. A.K. Seyfari, M. Bahadoran, P. Yupapin, *Nano Commun. Netw.* **29**, 100352 (2021)
31. M. Bahadoran, P. Yupapin, I.S. Amiri, *MiOTL* **63**(2), 653–661 (2021)
32. H. Su, X.G. Huang, *Sens. Actuators B Chem.* **126**(2), 579–582 (2007)
33. M. Bahadoran, A.F.A. Noorden, K. Chaudhary, M.S. Aziz, J. Ali, P. Yupapin, *Measurement* **58**, 215–220 (2014)
34. S.J. Mason, *Proc. IRE* **41**(9), 1144–1156 (1953)
35. S.J. Mason, *Proc. IRE* **44**(9), 920–926 (1956)
36. M. Bahadoran, P. Yupapin, *J. Theor. Appl. Phys.* **12**(2), 127–134 (2018)
37. M. Bahadoran, M. Aziz, A. Noorden, M. Jalil, J. Ali, P. Yupapin, *Dig. J. Nanomater. Biostruct.* **9**(3), 1095–1104 (2014)
38. T. Claes, W. Bogaerts, P. Bienstman, Presented at the SPIE NanoScience+ Engineering (2011) (unpublished)
39. C.-Y. Chao, L.J. Guo, *Appl. Phys. Lett.* **83**(8), 1527–1529 (2003)
40. M. Bahadoran, A.F.A. Noorden, F.S. Mohajer, M.H. Abd Mubin, K. Chaudhary, M.A. Jalil, J. Ali, P. Yupapin, *Artif. Cells Nanomed. Biotechnol.* **44**(1), 315–321 (2016)
41. Y. Sun, X. Fan, *Anal. Bioanal. Chem.* **399**(1), 205–211 (2011)
42. H. Zhu, I.M. White, J.D. Suter, P.S. Dale, X. Fan, *Opt. Express* **15**(15), 9139–9146 (2007)
43. L. Jin, M. Li, J.-J. He, *Opt. Commun.* **284**(1), 156–159 (2011)
44. M. Bahadoran, J. Ali, P.P. Yupapin, *Appl. Opt.* **52**(12), 2866–2873 (2013)
45. A.F.A. Noorden, A. Mohamad, M. Bahadoran, K. Chaudhary, M. Aziz, M.A. Jalil, J. Ali, P. Yupapin, *Jurnal Teknologi* **74**(8) (2015)
46. T. Claes, W. Bogaerts, P. Bienstman, Presented at the Biosensing and Nanomedicine IV (2011) (unpublished)
47. M. Bahadoran, P. Yupapin, *Eur. Phys. J. Plus* **133**(11), 487 (2018)
48. M. Aziz, S. Daud, M. Bahadoran, J. Ali, P.P. Yupapin, *J. Nonlinear Opt. Phys. Mater.* **21** (04) (2012)
49. T. Claes, J.G. Molera, K. De Vos, E. Schachtb, R. Baets, P. Bienstman, *Photonics J. IEEE* **1**(3), 197–204 (2009)
50. K. De Vos, I. Bartolozzi, E. Schacht, P. Bienstman, R. Baets, *Opt. Express* **15**(12), 7610–7615 (2007)
51. C.L. Arce, K. De Vos, T. Claes, K. Komorowska, D. Van Thourhout, P. Bienstman, *Photonics Technol. Lett. IEEE* **23**(13), 890–892 (2011)
52. M. Butt, S. Degtyarev, S. Khonina, Presented at the Journal of Physics: Conference Series (2018) (unpublished)
53. M. Vakili, M. Noori, *OQE* **51**(3), 1–9 (2019)
54. G.M. Hale, M.R. Querry, *Appl. Opt.* **12**(3), 555–563 (1973)
55. A. Kalate Seyfari, A.H. Zareian, M. Bahadoran, *J. Res. Many-body Syst.* (2021)

Chapter 13

Natural Radioactivity, Transfer Factor and Associated Radiological Risk in Commercially Cultivated Yam (*Dioscorea Rotundata*) in Northcentral Nigeria



Matthew Tikpangi Kolo, Oyeleke Ismail Olarinoye, Simon Olonkwoh Salihu, Hauwau Kulu Shuaibu, and Funmilayo Ayedun

Abstract Human food chain can become contaminated either by direct radionuclide deposition, absorption from radionuclide-polluted soil and water by plant roots and direct ingestion of polluted plants, soil or water by animals. In this study, activity concentrations of primordial radionuclides in soil and yam (*Dioscorea rotundata*) samples from a commercially cultivated yam farm in northcentral Nigeria were analyzed using a 3" × 3" NaI(Tl) gamma detector. Results show that mean specific activities of ^{238}U , ^{232}Th and ^{40}K in soil and yam samples were 40.36 ± 3.97 , 14.71 ± 0.80 , 385.63 ± 16.54 Bq kg $^{-1}$, and 31.11 ± 4.00 , 11.82 ± 0.72 , 466.96 ± 27.20 Bq kg $^{-1}$ respectively, which are within limits of safety set by the United Nations Scientific Committee on the Effect of Atomic Radiation. The average absorbed dose for soil samples was 43.63 nGy h $^{-1}$ with corresponding mean annual effective dose of 0.05 mSv y $^{-1}$. Yam samples recorded mean absorbed dose rate of 42.61 nGy h $^{-1}$ with corresponding mean annual effective dose of 0.05 mSv y $^{-1}$, which were within international safety limits. Computed average soil-to-yam transfer factor was 0.70, 0.83 and 1.23 respectively for ^{238}U , ^{232}Th and ^{40}K . Transfer factors for ^{238}U and ^{232}Th

M. T. Kolo (✉) · O. I. Olarinoye
Department of Physics, Federal University of Technology, Minna, Niger Sate, Nigeria
e-mail: matthewkolo@futminna.edu.ng

O. I. Olarinoye
e-mail: Leke.olarinoye@futminna.edu.ng

S. O. Salihu
Department of Chemistry, Federal University of Technology, Minna, Niger Sate, Nigeria

H. K. Shuaibu
Department of Physics, Nigerian Defence Academy, Kaduna, Kaduna State, Nigeria

F. Ayedun
Department of Physics, National Open University of Nigeria, Abuja, Nigeria
e-mail: fayedun@noun.edu.ng

were below unity, while for ^{40}K was significantly moderate, showing that bioaccumulation of natural radionuclides in the Nigerian grown yam does not pose any immediate radiological threat for public consumption. The yam tubers are therefore fit, not just for consumption, but also for export to other nations from a radiological perspective. Routine radiological checks of food crops are however encouraged in compliance with the ALARA provisions.

13.1 Introduction

Radiological contamination of dietary pathways by naturally occurring radionuclides has attracted great attention from a radiation protection perspective [1, 2]. The human food chain can become radiologically contaminated either by direct radionuclide deposition on plant leaves, uptake from contaminated soil and water by plant roots, and direct ingestion of contaminated plants, soil, water and animals [3]. Additional contamination pathway is fertilizer application to cultivated farms. Fertilizers, especially the phosphate based ones, are known to contain appreciable concentration of naturally occurring radioactive materials (^{238}U , ^{232}Th , and ^{40}K) and their radioactive daughters. Thus, extensive use of fertilizers in the agriculture sector for improved nutrient supply and increase crop yield can increase the amount of radionuclides in soils and consequential ingestion by humans through the food chain. One of the most potent exposure routes by which radionuclides get into the human body is by direct and/or indirect ingestion of contaminated agricultural and livestock products [4]. Radionuclides that enter the human body contribute essentially to the total radiation dose to different human body organs and constitute a long-term challenge to human health and safety [5, 6]. It is therefore important that the contamination levels of agricultural soils and agricultural products be investigated and the rate of radionuclide uptake be assessed via the transfer factor.

A very essential index for accurately predicting radionuclide migration from soil to plants and adequately estimating radionuclide concentrations in plants is the soil-to-plant transfer factor (TF) [7, 8]. TF is the proportion of radionuclide activity in a given plant to its corresponding activity in soil. It is one of the essential parameters in assessing the internal dose to humans via the ingestion route. TF is a critical parameter in the study of the impact of radionuclide releases in the human environment. Bioaccumulation of radionuclides in soils and their subsequent migration to plants are greatly affected by the nature of vegetation, soil types and soil pH, climatic conditions and solid/liquid distribution coefficient [9, 10]. It is therefore expedient to constantly undertake a localized assessment of TFs to estimate dose impact on the human population and assure food safety from the perspective of radiation protection. It will also help in modelling and predicting the future accumulation of primordial radionuclides in locally cultivated food crops. Several studies have been conducted in many countries to evaluate soil-to-plant transfer factors (TF) of natural and artificial radionuclides for most staple food crops. This study is one of the few that has been conducted in Nigeria.

Yam (*Dioscorea rotundata*) is one of the most staple food in Nigeria. North central Nigeria in particular is a region that is known for the cultivation and consumption of yam tubers. With the current drive by the Nigerian government towards making the agricultural sector the sole alternative economic base, Nigeria is now targeting huge foreign exchange earnings from the export of this commodity. This investigation has therefore become necessary to ensure that all agricultural products exported out of Nigeria meet international safety standards from a radiation protection perspective.

According to Jibiri et al. [5], “the three principal objectives pursued by the United Nations for sustainable food security include (i) ensuring access of all people to sufficient, nutritionally adequate and safe food; (ii) continued and sustainable contribution of agriculture to economic and social progress, and (iii) conservation and sustainable utilization of natural resources, including land, water, and genetic resource base for food and agriculture”. In compliance with the protocols for sustainable food security of the United Nations [11], therefore, it is paramount that radiological safety of all foodstuffs meant for human consumption at all levels is not compromised. This study therefore is aimed at determining the activity concentrations of ^{238}U , ^{232}Th and ^{40}K in Nigerian yam tubers and to evaluate the bioaccumulation status of these radionuclides. Results from this research will help in setting up radiometric control standards that will keep the effective doses due to ingestion of Nigerian foodstuffs as low as reasonably achievable.

13.2 Materials and Method

13.2.1 Sample Collection

Fifteen (15) commercially cultivated yam (*D. rotundata*) samples and their corresponding soil samples were collected randomly throughout the farm for analysis. This ensures that each yam plant and soil has equal probability of being sampled thus eliminating biased representation. At every sample collection point, the soil samples were cleaned of every contaminant like shrubs and pebbles. The samples were thereafter packaged into well labelled polyethylene bags for proper identification and preservation. Similarly, yam samples were thoroughly cleaned and sun-dried for about 3 h to eliminate surface moisture after which they were packed in clean, well-labelled bags for proper identification. Both the yam and the corresponding soil samples were finally transported to the laboratory for further preparation.

13.2.2 Sample Preparation

Soil samples were dried openly at room temperature for 72 h in the laboratory until a constant weight was achieved; indicating complete elimination of moisture content.

The dry soil samples were grounded thoroughly into powder and sieved with 2 mm sieving mesh. 371.9 ± 0.2 g of each soil sample was packed into sample containers, correctly labeled and tightly closed to stop radon gas from leaking away, since its daughter product, ^{214}Bi which is used to evaluate the concentration of the parent nuclide (^{238}U), must be allowed to accumulate with time. The sealed samples were then stored for a period of thirty five (35) days to allow for secular equilibrium between the primordial radionuclides and their decay daughters [12, 13].

Likewise, yam samples were peeled. The edible portions were sliced into pieces, thoroughly washed with deionized water and dried at room temperature for 72 h to guarantee zero moisture content in the samples. The dry yam samples were grounded into powder, sieved with 2 mm mesh sieve and assiduously homologized. 180.8 ± 0.1 g of the homologized samples were filled into plastic sample containers which were labelled accurately, sealed at three levels and finally stored for thirty five (35) days for radiological equilibration before gamma spectrometric analysis.

13.2.3 Sample Analysis

Each sample (soil and yam) was radiometrically analysed using a $3'' \times 3''$ NaI(Tl) detector by Scintillation Technologies USA, at Ladoke Akintola University of Science and Technology Ogbomosho (LAUTECH), Nigeria. The detector which is coupled to a computer based multichannel analyzer (MCA) with ACCUSPEC computer program used for data acquisition and gamma spectra analysis is housed in a 6 cm thick lead shield to reduce the background radiation levels. The inside of the detector is also lined with cadmium and copper sheets to minimize spectrum interference by foreign frequencies. Efficiency and energy calibration of the detector was carried out before analysis using ^{137}Cs and ^{60}Co standard isotopic sources over the energy range of 200 keV to 2.810 MeV. The background was obtained by counting an empty sealed sample container similar to that of the samples for 36,000 s. Each sample was afterward counted for the same period as the empty container. Activity concentration of ^{238}U was determined from the 1764 keV gamma ray line of ^{214}Bi , while ^{232}Th activity was evaluated from 2614 keV gamma ray line of ^{208}Tl . Activity concentration of ^{40}K was obtained from its single gamma transition of 1460.822 keV. Activity concentrations, A, of ^{238}U , ^{232}Th and ^{40}K were calculated using the equation [14–16]:

$$A(\text{Bq kg}^{-1}) = \frac{C}{\varepsilon_{\gamma} \times I_{\gamma} \times M} \quad (13.1)$$

where C (counts per second) is the net count rate of each sample, ε_{γ} (E) is the detector photo-peak efficiency (%) at respective gamma-ray peak, I_{γ} is the corresponding gamma ray intensity and M is the sample mass in kg.

13.2.4 Computation of Radiation Hazard Indices

Hazard indices used in this study to quantify the radiation burden on the population as a result of exposure are as follows:

Radium Equivalent Activity (R_{eq}). Cumulative assessment of radiation hazard associated with ^{238}U , ^{232}Th and ^{40}K was done using radium equivalent activity. It's a single index computed based on the fact that same gamma dose is produced by 370 Bq kg^{-1} of ^{226}Ra , 259 Bq kg^{-1} of ^{232}Th and 4810 Bq kg^{-1} of ^{40}K ¹⁷⁻¹⁹. R_{eq} was calculated from the equation [12, 17]:

$$R_{eq} = A_U + 1.43A_{Th} + 0.077A_K \quad (13.2)$$

where A_U , A_{Th} and A_K are the activity concentrations of ^{238}U , ^{232}Th and ^{40}K respectively in Bq kg^{-1} .

Absorbed Dose Rate (D_R). Dose rate in air at 1 m above ground level due to gamma ray emissions from primordial radionuclides was computed using the equation [12, 18, 19]:

$$D_R (\text{nGy h}^{-1}) = 0.462A_U + 0.604A_{Th} + 0.0417A_K \quad (13.3)$$

where A_U , A_{Th} and A_K are the respective activity concentrations obtained from Eq. (13.1), 0.462, 0.604 and 0.0417 Bq kg^{-1} are conversion factors for ^{226}Ra , ^{232}Th and ^{40}K respectively provided by UNSCEAR [20].

Annual Effective Dose Equivalent (AEDE). The factor of 0.70 Sv Gy^{-1} which converts absorbed dose rate to effective dose, and the outdoor occupancy factor of 0.2, were used in computing the annual effective dose equivalent. It was assumed that about 20% of the total time of an individual is spent outdoor [12]. AEDE (measured in mSv y^{-1}) in outdoor air, was computed from Eq. (13.4).

$$\text{AEDE} (\text{mSv y}^{-1}) = D_R \times 1.21 \times 10^{-3} \quad (13.4)$$

External Hazard Index (H_{ex}). Radiation hazard acquired by an individual due to gamma dose exposure was evaluated using the external hazard index. H_{ex} was calculated from the equation [12, 18]:

$$H_{ex} = \frac{A_U}{370} + \frac{A_{Th}}{259} + \frac{A_K}{4810} \quad (13.5)$$

Internal Hazard Index (H_{in}). Internal hazard index give an indication of the likelihood of radiation incidence occurring in respiratory organs as a result of internal susceptibility to radon and its daughters. H_{in} was computed using the equation [12]:

$$H_{in} = \frac{A_U}{185} + \frac{A_{Th}}{259} + \frac{A_K}{4810} \quad (13.6)$$

where A_{Ra} , A_{Th} , and A_K are the activity concentrations of ^{226}Ra , ^{232}Th and ^{40}K respectively. UNSCEAR [12] provides the limit of unity for both H_{ex} and H_{in} below which any radiation incidence is inconsequential.

Excess Life Cancer Risk (ELCR). A measure of probability of occurrence of cancer incidence due to radiation exposure is expressed by the excess life cancer risk. ELCR was computed using the calculated AEDE from the equation [19, 21]:

$$\text{ELCR} = \text{AEDE} \times \text{DL} \times \text{RF} \quad (13.7)$$

where DL is the stipulated life span of 70 years and RF is risk factor of 0.05 Sv^{-1} for stochastic effects [22].

13.2.5 Transfer Factor (TF)

The rate of migration of radionuclides from soil to plants is evaluated using the transfer factor. TFs which were obtained from the activity concentrations of primordial radionuclides in both soil and yam samples were computed from the equation [23, 24]:

$$TF = \frac{\text{Activity concentration of a given radionuclide} (\text{Bq kg}^{-1} \text{ dry weight})}{\text{Activity concentration of the radionuclide in soil} (\text{Bq kg}^{-1} \text{ dry weight})} \quad (13.8)$$

For the computation of TFs in this research, average activity concentration in upper 20–25 cm top soil layer was considered. TF values less than unity show a low or inconsequential plant uptake of radionuclides from the soil, while values above unity give an indication of active radionuclide bioaccumulation.

13.3 Results and Discussion

Activity concentrations of primordial radionuclides in soil samples along with the corresponding radiation hazard indices are presented in Table 13.1.

Activity concentration of ^{238}U ranged from 8.39 ± 1.37 to $51.18 \pm 6.42 \text{ Bq kg}^{-1}$ with an average value of $40.36 \pm 3.97 \text{ Bq kg}^{-1}$, while that of ^{232}Th ranged from 8.39 ± 0.50 to $21.68 \pm 1.25 \text{ Bq kg}^{-1}$ with a mean value of $14.71 \pm 0.80 \text{ Bq kg}^{-1}$.

Table 13.1 Activity concentrations (Bq kg^{-1}) of ^{238}U , ^{232}Th and ^{40}K in soil samples

S. No	Sample ID	Activity concentrations (Bq kg^{-1})			
		^{238}U	^{232}Th	^{40}K	Ra_{eq}
1	SSS 01	14.26 ± 1.74	20.25 ± 1.20	383.75 ± 12.29	72.77
2	SSS 02	15.80 ± 2.26	13.69 ± 0.81	267.84 ± 14.07	56.00
3	SSS 03	70.55 ± 4.70	16.98 ± 0.53	436.93 ± 22.92	128.48
4	SSS 04	51.18 ± 6.42	17.61 ± 0.81	398.63 ± 16.20	107.06
5	SSS 05	32.50 ± 4.08	18.51 ± 1.10	395.79 ± 19.72	89.45
6	SSS 06	47.40 ± 5.88	9.89 ± 0.59	331.24 ± 17.36	87.05
7	SSS 07	20.54 ± 2.38	14.47 ± 0.63	408.06 ± 16.20	72.65
8	SSS 08	49.04 ± 6.15	11.73 ± 0.70	422.83 ± 22.20	98.37
9	SSS 09	48.22 ± 6.24	15.18 ± 0.91	355.90 ± 18.67	97.33
10	SSS 10	44.81 ± 4.41	21.69 ± 1.25	398.16 ± 5.10	106.49
11	SSS 11	bdl	9.05 ± 0.54	359.32 ± 8.47	–
12	SSS 12	88.73 ± 1.83	10.39 ± 0.50	360.00 ± 18.85	131.31
13	SSS 13	8.39 ± 1.27	16.79 ± 0.99	387.38 ± 20.26	62.23
14	SSS 14	bdl	14.62 ± 0.87	489.77 ± 15.28	–
15	SSS 15	33.24 ± 4.20	9.78 ± 0.58	388.86 ± 20.45	77.17
	Min	8.39 ± 1.27	8.39 ± 0.50	101.16 ± 5.10	56.00
	Max	51.18 ± 6.42	21.69 ± 1.25	436.93 ± 22.92	131.31
	Mean	40.36 ± 3.97	14.71 ± 0.80	385.63 ± 16.54	91.26
World average (UNSCEAR, 2000)		35	30	400	≤ 370

^{40}K activity values ranged from 101.16 ± 5.10 to 436.83 ± 22.92 Bq kg^{-1} with an average of 385.63 ± 16.54 Bq kg^{-1} . These values are within the world average values of 35, 30 and 400 Bq kg^{-1} for ^{238}U , ^{232}Th and ^{40}K respectively [12]. Computed radium equivalent activity (Ra_{eq}) varied from 56.00 Bq kg^{-1} to 131.31 Bq kg^{-1} with a mean value of 91.26 Bq kg^{-1} . The natural distribution pattern of ^{238}U , ^{232}Th and ^{40}K in the soil samples is shown in Fig. 13.1.

Radiation hazard indices were computed using Eqs. (13.3)–(13.7) and the results are presented in Table 13.2. Calculated mean value of absorbed dose (D_R) at 1 m was 42.63 nGy h^{-1} with a corresponding annual effective dose (AEDE) of 0.05 mSv y^{-1} , while computed external and internal hazard indices were below unity (Table 13.1). Calculated excess lifetime cancer risk (ELCR) varied from 0.11×10^{-3} to 0.26×10^{-3} with a mean value of 0.18×10^{-3} . All computed radiation hazard indices were below their respective world average values of 370 Bq kg^{-1} , 58 nGy h^{-1} , 0.07 mSv y^{-1} and 0.18×10^{-3} documented in the UNSCEAR report [12].

Table 13.3 shows the specific activities of ^{238}U , ^{232}Th and ^{40}K in the commercial yam (*D. rotundata*) samples together with the corresponding hazard indices. Specific activities of ^{238}U , ^{232}Th and ^{40}K ranged from 2.88 ± 0.44 to 76.68 ± 9.73 Bq kg^{-1} ,

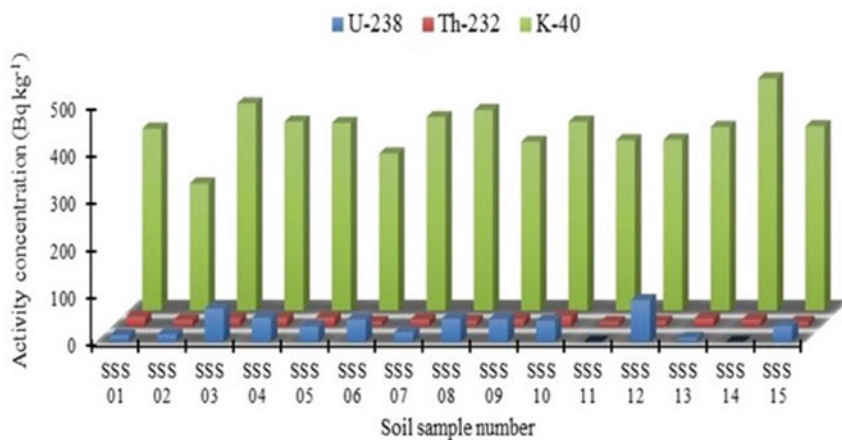


Fig. 13.1 Activity concentrations of ^{238}U , ^{232}Th and ^{40}K in the soil samples

Table 13.2 Radiation hazard indices for the soil samples

S. No	Sample ID	D_R (nGy h ⁻¹)	AEDE (mSv y ⁻¹)	H_{ex}	H_{in}	ELCR ($\times 10^{-3}$)
1	SSS 01	34.82	0.04	0.20	0.24	0.15
2	SSS 02	26.74	0.03	0.15	0.19	0.11
3	SSS 03	61.07	0.07	0.35	0.54	0.26
4	SSS 04	50.90	0.06	0.29	0.43	0.22
5	SSS 05	42.70	0.05	0.24	0.33	0.18
6	SSS 06	41.69	0.05	0.24	0.36	0.18
7	SSS 07	35.25	0.04	0.20	0.25	0.15
8	SSS 08	47.37	0.06	0.27	0.40	0.20
9	SSS 09	46.29	0.06	0.26	0.39	0.20
10	SSS 10	50.41	0.06	0.29	0.41	0.21
11	SSS 11	–	–	–	–	–
12	SSS 12	62.28	0.08	0.35	0.59	0.26
13	SSS 13	30.17	0.04	0.17	0.19	0.13
14	SSS 14	–	–	–	–	–
15	SSS 15	37.48	0.05	0.21	0.30	0.16 ara>
	Min	26.74	0.03	0.15	0.19	0.11
	Max	62.28	0.08	0.35	0.59	0.26
	Mean	43.63	0.05	0.25	0.36	0.18
World average (UNSCEAR, 2000)		[60 (18-93)]	1.00	≤ 1	≤ 1	0.29

Table 13.3 Activity concentrations (Bq kg^{-1}) of ^{238}U , ^{232}Th and ^{40}K in yam samples

S. No	Sample ID	Activity concentrations (Bq kg^{-1})			
		^{238}U	^{232}Th	^{40}K	Ra_{eq}
1	DRD 01	11.77 ± 1.64	7.73 ± 0.46	438.31 ± 38.37	56.57
2	DRD 02	bdl	10.05 ± 0.78	499.31 ± 26.14	–
3	DRD 03	72.24 ± 8.49	14.06 ± 0.84	627.94 ± 32.77	140.70
4	DRD 04	33.24 ± 4.67	19.77 ± 1.30	530.22 ± 28.00	102.34
5	DRD 05	12.75 ± 1.74	14.31 ± 0.85	484.54 ± 30.41	70.52
6	DRD 06	37.19 ± 4.82	8.13 ± 0.49	460.56 ± 23.99	84.28
7	DRD 07	29.57 ± 2.93	15.43 ± 0.83	502.08 ± 28.91	90.30
8	DRD 08	25.92 ± 3.73	bdl	502.49 ± 26.35	–
9	DRD 09	2.88 ± 0.44	10.42 ± 0.63	435.56 ± 22.84	51.32
10	DRD 10	bdl	13.02 ± 0.78	312.50 ± 26.83	–
11	DRD 11	bdl	7.48 ± 0.45	408.86 ± 21.43	–
12	DRD 12	76.68 ± 9.73	9.69 ± 0.58	346.36 ± 20.73	117.21
13	DRD 13	bdl	bdl	446.36 ± 28.50	–
14	DRD 14	24.85 ± 3.28	9.88 ± 0.59	514.42 ± 26.85	78.59
15	DRD 15	15.14 ± 2.08	13.64 ± 0.81	494.88 ± 25.87	72.75
	Min	2.88 ± 0.44	7.48 ± 0.45	312.50 ± 20.73	51.32
	Max	76.68 ± 9.73	19.77 ± 1.30	627.94 ± 38.37	140.70
	Mean	31.11 ± 4.0	11.82 ± 0.72	466.96 ± 27.20	86.46

7.48 ± 0.45 to $19.77 \pm 1.30 \text{ Bq kg}^{-1}$ and 312.50 ± 20.73 to $627.94 \pm 38.37 \text{ Bq kg}^{-1}$ respectively, with their corresponding average values of $31.11 \pm 4.0 \text{ Bq kg}^{-1}$, $11.82 \pm 0.72 \text{ Bq kg}^{-1}$ and $466.96 \pm 27.20 \text{ Bq kg}^{-1}$ in sequence. Calculated values for radium equivalent activity (Ra_{eq}) varied from 51.32 Bq kg^{-1} to $140.70 \text{ Bq kg}^{-1}$, with an average value of 86.46 Bq kg^{-1} . Computed averages for ^{238}U and ^{232}Th are within the international safety limits documented in UNSCEAR [12] report. Mean specific activity value for ^{40}K in yam samples is however slightly higher which may not pose any radiological threat since ^{40}K is an essential element needed for body development. Results of this study fall within the same range as that of similar studies in other parts of Nigeria and in Ghana [2, 3, 9, 25, 26]. Measured mean specific activities of ^{238}U , ^{232}Th and ^{40}K in this study were however found to be lower than those obtained for yam tubers in Jos, northcentral Nigeria, which were documented to be $684.5 \pm 40.6 \text{ Bq kg}^{-1}$ for ^{40}K , $85.5 \pm 10.2 \text{ Bq kg}^{-1}$ for ^{238}U and $89.8 \pm 6.2 \text{ Bq kg}^{-1}$ for ^{232}Th [5]. These enhanced radioactivity contents are due to continuous tin mining activities that have been going on in Jos, northcentral Nigeria for decades.

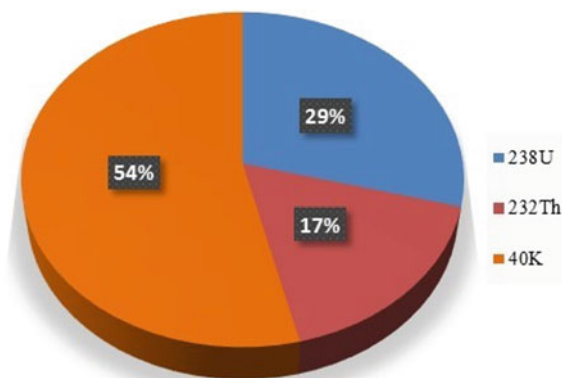
Results of the calculated radiation hazard indices for yam samples are presented in Table 13.4. Average values of 42.21 nGy h^{-1} , 0.05 mSv y^{-1} and 0.32 were recorded for the absorbed dose (D_{R}), annual effective dose (AEDE) and internal hazard index

Table 13.4 Radiation hazard indices for yam samples

S. No	Sample ID	D_R (nGy h ⁻¹)	AEDE (mSv y ⁻¹)	$H_{in} \leq 1$	ELCR ($\times 10^{-3}$)
1	DRD 01	28.38	0.03	0.18	0.12
2	DRD 02	–	–	–	–
3	DRD 03	68.05	0.08	0.58	0.29
4	DRD 04	49.41	0.06	0.37	0.21
5	DRD 05	34.74	0.04	0.22	0.15
6	DRD 06	41.30	0.05	0.33	0.17
7	DRD 07	43.92	0.05	0.32	0.19
8	DRD 08	–	–	–	–
9	DRD 09	25.79	0.03	0.15	0.11
10	DRD 10	–	–	–	–
11	DRD 11	–	–	–	–
12	DRD 12	55.72	0.07	0.52	0.24
13	DRD 13	–	–	–	–
14	DRD 14	38.90	0.05	0.28	0.16
15	DRD 15	35.87	0.04	0.24	0.15
	Min	25.79	0.03	0.15	0.11
	Max	68.05	0.08	0.58	0.29
	Mean	42.21	0.05	0.32	0.18

(H_{in}) respectively. These values were found to be within the internationally acceptable standards as prescribed by the UNSCEAR [12]. The percentage contribution of primordial radioactive nuclides to the total absorbed dose is shown in Fig. 13.2. Computed excess lifetime cancer risk (ELCR) ranged between 0.11×10^{-3} and 0.29×10^{-3} with a mean of 0.18×10^{-3} . This result is lower than the safety limit of 0.29×10^{-3} documented by UNSCEAR [12] and 0.05 stipulated in the ICRP reports

Fig. 13.2 Percentage contributions of ^{238}U , ^{232}Th and ^{40}K to total absorbed dose



for low-level radiations. This shows that the probability of cancer incidences due to consumption of Nigerian grown yam tubers is insignificant.

Soil-to-yam transfer factors (TF) for ^{238}U , ^{232}Th and ^{40}K for the commercial yam tubers were computed using Eq. (13.8) and the results are presented in Table 13.5. Root uptake has been identified as the principal process by which radionuclides can accumulate in plants. As seen in Table 13.5, TF for ^{238}U ranged between 0.06 and 1.44, with an average value of 0.70. TF for ^{232}Th varied from 0.38 to 1.39, with a mean value of 0.83, while TF for ^{40}K has the highest value of 1.86 and a mean of 1.23. Rate of radionuclide transfer from soil to yam shows a decreasing trend of ^{40}K end of ^{40}K (1.23) > ^{232}Th (0.83) > ^{238}U (0.70).

Mean TF values for ^{238}U , ^{232}Th and ^{40}K in this study compare favourably with those obtained in similar studies [3, 9, 27, 28]. Computed TFs in this study are below the recommended value of 2.7 for tubers [4]. Thus, the rate of radionuclides transfer and their subsequent accumulation in Nigerian yam tubers are moderate and do not constitute any immediate radiological incidence to the consumers.

Table 13.5 Soil-to-yam transfer factors (TF) for ^{238}U , ^{232}Th and ^{40}K in yam samples

S. No	Soil sample	Yam sample	^{238}U	^{232}Th	^{40}K
1	SSS 01	DRD 01	0.83	0.38	1.14
2	SSS 02	DRD 02	–	0.73	1.86
3	SSS 03	DRD 03	1.02	0.83	1.44
4	SSS 04	DRD 04	0.65	1.12	1.33
5	SSS 05	DRD 05	0.39	0.77	1.22
6	SSS 06	DRD 06	0.78	0.82	1.39
7	SSS 07	DRD 07	1.44	1.07	1.23
8	SSS 08	DRD 08	0.53	–	1.19
9	SSS 09	DRD 09	0.06	0.69	1.22
10	SSS 10	DRD 10	–	0.60	0.78
11	SSS 11	DRD 11	–	0.83	1.14
12	SSS 12	DRD 12	0.86	0.93	0.96
13	SSS 13	DRD 13	–	–	1.15
14	SSS 14	DRD 14	–	0.68	1.05
15	SSS 15	DRD 15	0.46	1.39	1.27
	Min		0.06	0.38	0.78
	Max		1.44	1.39	1.86
	Mean		0.70	0.83	1.23

13.4 Conclusion

Activity concentrations of ^{238}U , ^{232}Th and ^{40}K in commercially cultivated yam tubers and the corresponding cultivated soil samples were measured using gamma spectrometric technique which employs $3'' \times 3''$ NaI(Tl) gamma detector. Mean specific activities of ^{238}U , ^{232}Th and ^{40}K in yam tubers and soil samples are within their respective world average values presented in the United Nations Scientific Committee on the Effects of Atomic Radiation reports. Computed radiation hazard indices which are used in estimating any possible radiological health risk in the studied samples and the subsequent radiation dose associated with yam tuber consumption are below their permissible limits. Soil to plant transfer factor (TF) for the primordial radionuclides computed for the studied yam samples appears to be generally moderate. This indicates that activity concentrations of natural radionuclides in the Nigerian grown yam tubers do not pose any immediate radiological threat for public consumption and hence are fit for export to other nations from a radiological perspective. Continuous radiological monitoring of the food crop is however encouraged to check accumulation effects due to long-term consumption.

Acknowledgements The authors express their deep gratitude to the management and staff Ladoke Akintola University of Science and Technology Ogbomosho (LAUTECH), Nigeria.

References

1. S. Tang, Z. Chen, H. Li, J. Zheng, *Environ. Pollut.* **125**(3), 305–312 (2003)
2. A. Arogunjo, E. Ofuga, M. Afolabi, *J. Environ. Radioact.* **82**(1), 1–6 (2005)
3. A.O. Gregory, E. Agbalagba, *Environ. Earth Sci.* **71**(4), 1541–1549 (2014)
4. IAEA, International Atomic Energy Agency, Vienna (2010)
5. N. Jibiri, I. Farai, S. Alausa, *Radiat. Environ. Biophys.* **46**(1), 53–59 (2007)
6. N. Adesiji, J. Ademola, *Niger. J. Pure Appl. Phys.* **9**(1), 6–10 (2019)
7. S. Uchida, K. Tagami, I. Hirai, *J. Nucl. Sci. Technol.* **44**(5), 779–790 (2007)
8. S.R. Chakraborty, R. Azim, A.R. Rahman, R. Sarker, *J. Phys. Sci.* **24**(1), 95 (2013)
9. I.N.Y. Doyi, D.K. Essumang, A.K. Agyapong, S. Asumadu-Sarkodie, *J. Environ. Radioact.* **182**, 138–141 (2018)
10. H.D. Van, T.D. Nguyen, A. Peka, M. Hegedus, A. Csordas, T. Kovacs, *J. Environ. Radioact.* **223**, 106416 (2020)
11. R. Pérez-Escamilla, *Curr. Dev. Nutr.* **1**(7), e000513 (2017)
12. UNSCEAR (2000)
13. K. Asaduzzaman, M.U. Khandaker, Y.M. Amin, D.A. Bradley, R.H. Mahat, R.M. Nor, *J. Environ. Radioact.* **135**, 120–127 (2014)
14. M.U. Khandaker, P. Jojo, H. Kassim, Y. Amin, *Radiat. Prot. Dosimet.* **152**(1–3), 33–37 (2012)
15. M. Kolo, S. Aziz, M. Khandaker, K. Asaduzzaman, Y. Amin, *Environ. Sci. Pollut. Res.* **22**(17), 13127–13136 (2015)
16. M.T. Kolo, M. Gomina, B. Awojoyogbe, O. Olarinoye, *J. Nucl. Technol. Appl. Sci.* **8**(1), 97–111 (2020)
17. J. Beretka, P. Mathew, *Health Phys.* **48**(1), 87–95 (1985)
18. U. Cevik, H. Baltas, A. Tabak, N. Damla, *J. Hazard. Mater.* **182**(1), 531–535 (2010)

19. R. Ravisankar, K. Vanasundari, M. Suganya, Y. Raghu, A. Rajalakshmi, A. Chandrasekaran, S. Sivakumar, J. Chandramohan, P. Vijayagopal, B. Venkatraman, *Appl. Radiat. Isot.* **85**, 114–127 (2014)
20. UNSCEAR, *Effects of Ionizing Radiation: Report to the General Assembly, with Scientific Annexes*. (United Nations Publications, 2008)
21. H. Taskin, M. Karavus, P. Ay, A. Topuzoglu, S. Hidiroglu, G. Karahan, *J. Environ. Radioact.* **100**(1), 49–53 (2009)
22. ICRP Publication 60: 1990 Recommendations of the International Commission on Radiological Protection. (Elsevier Health Sciences, 1991)
23. H.H. Azeez, H.H. Mansour, S.T. Ahmad, *Appl. Radiat. Isot.* **147**, 152–158 (2019)
24. S.B. Ibikunle, A.M. Arogunjo, O.S. Ajayi, *Sci. Afr.* **3**, e00062 (2019)
25. G. Avwiri, A. Alao, *Int. J. Sci. Res.* **2**(1), 468–473 (2013)
26. N.N. Jibiri, B.C. Eke, *Int. J. Phys. Res. Appl.* **4**, 6–14 (2021)
27. IAEA (2009)
28. A.K. Ademola, *Radiat. Prot. Environ.* **42**(3), 112 (2019)

Chapter 14

Analysis of Electrical Transport Properties in $\text{Nd}_{0.75}\text{Na}_{0.25-x}\text{K}_x\text{MnO}_3$ Manganites



S. Shamsuddin, N. Ibrahim, M. Z. H. Mayzan, Fahmiruddin Esa, and S. A. Razali

Abstract K doping in the compound of $\text{Nd}_{0.75}\text{Na}_{0.25-x}\text{K}_x\text{MnO}_3$ ($x = 0, 0.05$ and 0.10) manganites have been investigated to study its effect on electrical transport properties as well as electrical behavior analysis. DC electrical resistivity measurement showed all samples exhibit insulating behavior from room temperature down to 50 K. However, analysis of $\ln\rho/dT^{-1}$ versus T revealed a signal which observed at temperature 210 K for $x = 0$ sample. Interestingly the peaks were then shifted to the lower temperature around 190 K and 165 K for $x = 0.05$ and $x = 0.1$ samples respectively, indicate the existence of charge ordering state in the compound. Apart from that, the electrical analysis suggested that the variable range hopping (VRH) model was selected as the best model to explain the electrical behavior of the samples as the value of R^2 approached to 99.99% compared to the small polaron hopping (SPH) model.

14.1 Introduction

Hole-doped perovskite with general formula $\text{R}_{1-x}\text{A}_x\text{MnO}_3$, where R is trivalent rare earth ion (Pr^{3+} , La^{3+} , Nd^{3+}) and A is alkaline monovalent (Na^+ , K^+ , Ag^+) or divalent (Ca^{2+} , Ba^{2+} , Sr^{2+}) ion, possess extraordinary properties such as colossal magnetoresistance (CMR), metal–insulator (MI) transition, ferromagnetic–paramagnetic

S. Shamsuddin (✉) · M. Z. H. Mayzan · F. Esa · S. A. Razali
Ceramic and Amorphous Group, Faculty of Applied Sciences and Technology, Pagoh Higher Education Hub, Universiti Tun Hussein Onn Malaysia, 84600 Panchor, Johor, Malaysia
e-mail: suhadir@uthm.edu.my

M. Z. H. Mayzan
e-mail: zulhilmi@uthm.edu.my

F. Esa
e-mail: fahmir@uthm.edu.my

N. Ibrahim
Faculty of Applied Sciences, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia
e-mail: noraz954@salam.uitm.edu.my

behavior as well as their potential application in magnetic sensor [1–6]. The occurrence of ferromagnetic-metallic transition is explained by double exchange mechanism, which involves the electron hopping process between pairs of Mn^{3+} and Mn^{4+} ion [7]. However, insulating behavior and revival of metallic behavior in excessive magnetic ion doped was not sufficiently enough to explained others mechanism such as Jahn–Teller (JT) effect [8], electron–phonon coupling and charge ordering (CO) [9]. In particular, the $\text{Nd}_{0.75}\text{Na}_{0.25}\text{MnO}_3$ compound has received fair notable attention as it exhibits a CO transition at relatively high temperature, $T_{\text{CO}} \sim 170$ K compared to AFM interaction [10, 11].

The percentage amount of alkaline metal ions in manganites are said to have a great influence on the magnetic and transport properties of perovskite material [12]. Previous study stated that substitution of K^+ in $\text{Pr}_{0.75}\text{Na}_{0.25-x}\text{K}_x\text{MnO}_3$ has been weakened the CO state and actuated the ferromagnetic–paramagnetic transition [2]. This statement supported by Thaljaoui et al., stated that when potassium substitution content increases, it has been shifted paramagnetic to the ferromagnetic transition temperature [13]. However, most of the substitution does not show full suppression of metallic state most probably due to dopant level. As such, the influence of magnetic ions on host composition needs to be further investigated to understand the role of magnetic ions in the CO phase. Hence, electrical resistivity measured at low temperature would be very interesting to study as this could reveal the information of CO state in the sample as well as to provide a better understanding on the electrical transport properties.

On the other hand, the electrical analysis also plays a crucial role for explaining the variation of electrical resistivity in the material that describing the electron scattering model and hopping model for metallic and insulating part respectively. Research proved that resistivity behavior in some manganites is suggested as a combination of weak electron localization and some scattering mechanisms, such as electron–electron, electron–magnon and electron–phonon scattering processes in ferromagnetic phase [14]. Alternatively, to understand the electrical transport properties at the insulating region, the resistivity data were fitted using two different hopping models, namely, variable range hopping (VRH) and small polaron hopping (SPH). VRH models describe the electrical analysis in manganite based on doping-induced disorder caused by the random distribution of defects, while SPH model implies the existence of strong electron–phonon coupling caused by the Jahn–Teller lattice distortion [2]. However, to the best of our knowledge, study on reporting the electrical resistivity data with hopping model using $\text{Nd}_{0.75}\text{Na}_{0.25-x}\text{K}_x\text{MnO}_3$ ($0 \leq x \leq 0.10$) manganite has not been reported previously. Thus, analyze of the electrical resistivity data will give a better understanding of the hopping process in this compound.

This study reports the DC electrical resistivity measurement on $\text{Nd}_{0.75}\text{Na}_{0.25-x}\text{K}_x\text{MnO}_3$ ($0 \leq x \leq 0.10$) manganite with the aim to investigate the effect of partial K-doped on electrical transport and analyze the electrical resistivity data as well as provide a better understanding of CO in the material.

14.2 Experiment Methods

The standard conventional solid state reaction method was used to synthesis $\text{Nd}_{0.75}\text{Na}_{0.25-x}\text{K}_x\text{MnO}_3$ ($x = 0, 0.05$ and 0.10) manganite samples by mixing high purity ($\geq 99.99\%$) powder of stoichio-metric quantities Nd_2O_3 , Na_2CO_3 , K_2CO_3 and MnO_2 in ambient atmosphere. The powders were carefully ground and calcined in air at 1000°C for 24 hour with several intermediate grindings. The mixture powders then compacted into a pellet form with a dimensional of 13 mm diameter and 2–3 mm thickness. Finally, the samples in pellet form undergo sintering process at 1200°C for 24 hour in air with heating rate of $15^\circ\text{C}/\text{min}$ and then slowly cooled at room temperature at the rate of $1^\circ\text{C}/\text{min}$ to ensure better crystallization. All the finishing samples were then proceeded for sample characterization with powder XRD measurement with $\text{CuK}\alpha$ radiation (1.5440 \AA) at scanning rate $2^\circ\text{C}/\text{min}$ using XRD Bruker D8 Advance model in room temperature condition. XRD analysis revealed all samples were indexed as an orthorhombic perovskite-structured where whose unit cell belongs to the space group $Pnma$ reliable with the previous study [11, 15]. Investigation on transport properties were performed with silver paste contact in a Janis Cryostat Model CCS-350ST using four-point-probe method at low temperature condition. Electrical analyses were carried out using the VRH and SPH models to propose a suitable model to elucidate the electrical resistivity data.

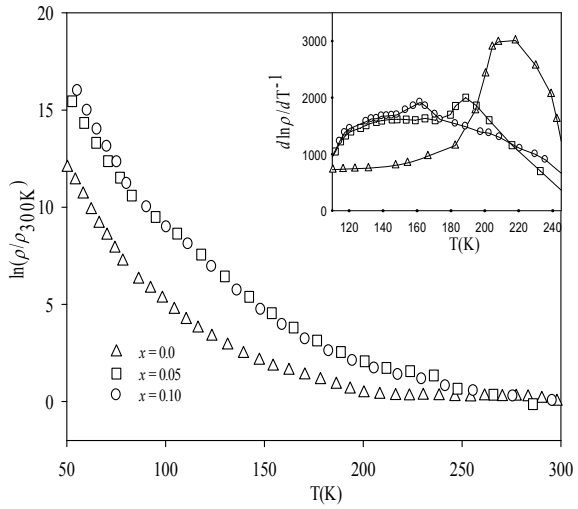
14.3 Results and Discussion

The temperature dependence of electrical resistivity for the $x = 0, 0.05$ and 0.1 samples were shown in Fig. 14.1. From the result, all samples showed insulating transport behavior in the temperature range of 50–300 K. However, it shown that from the $d\ln\rho/dT^{-1}$ versus T curve (inset Fig. 14.1), a peak around 210 K and 190 K have been observed for $x = 0$ and 0.05 samples respectively while a broader peak was observed for the $x = 0.10$ sample around 165 K. Coincidentally, it was found the location of the peaks was in the range of the charge ordering transition temperature, T_{CO} as reported by Zawawi et al. [15] and Fan et al. [16]. In addition, the shifting of the peaks from 210 K ($x = 0$) to 165 K ($x = 0.10$) samples indicate a weakening of CO state as a result of increasing of K content in the compound.

In order to understand the electrical behavior of monovalent doped in $\text{Nd}_{0.75}\text{Na}_{0.25-x}\text{K}_x\text{MnO}_3$ ($0 \leq x \leq 0.10$) manganites, the experimental resistivity data were fitted using two different hopping models, namely, VRH and SPH models [17–19]. For the SPH model, the electrical resistivity was fitted using Eq. 14.1 [1, 2, 14].

$$\rho = BT \exp(E_a/k_B T) \quad (14.1)$$

Fig. 14.1 Temperature dependence of electrical resistivity for $\text{Nd}_{0.75}\text{Na}_{0.25-x}\text{K}_x\text{MnO}_3$ ($0 \leq x \leq 0.10$). Inset indicates the $d \ln \rho / dT^{-1}$ versus T for $\text{Nd}_{0.75}\text{Na}_{0.25-x}\text{K}_x\text{MnO}_3$

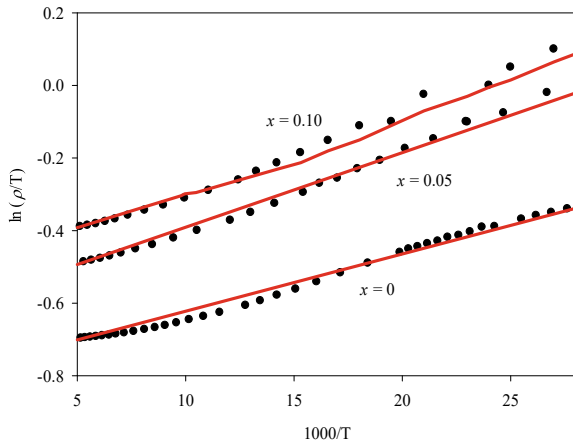


where E_a is the activation energy and B is the resistivity coefficient. E_a values for all samples were calculated from the gradient of $\ln \rho / T$ versus $1000/T$ graph.

Figure 14.2 showed the plotted SPH model for $\text{Nd}_{0.75}\text{Na}_{0.25-x}\text{K}_x\text{MnO}_3$ ($0 \leq x \leq 0.10$) and solid lines represented the fitting data. For VRH model, the resistivity data were fitted using following Eq. 14.2 [20, 21]

$$\rho = \rho_{0m} \exp(T_{0m}/T)^{1/4} \tag{14.2}$$

Fig. 14.2 Plot of $\ln \rho / T$ versus $1000/T$ for $\text{Nd}_{0.75}\text{Na}_{0.25-x}\text{K}_x\text{MnO}_3$ ($x = 0-0.10$) samples. The solid line indicates the best-fit to the SPH model



where ρ_{0m} is the Mott residual resistivity and T_{0m} is the Mott characteristic temperature and is related to the carrier localization length, which can be determined from the slope of $\ln(\rho)$ versus $T^{-1/4}$ plot.

From the obtained values of T_{0m} , the density states at the Fermi level, $N(E_F)$, hopping energy at room temperature, E_h (300 K) and mean hopping distance, R_h , were calculated using the following Eqs. 14.3, 14.4 and 14.5.

$$E_h(T) = 1/4k_B(T)^{3/4}(T_{0m})^{1/4} \tag{14.3}$$

$$N(E_F) = 18a^3/(k_B T_{0m}) \tag{14.4}$$

$$R_h(T) = \left(\frac{3}{8}\right)a(T_{0m}/T)^{1/4} \tag{14.5}$$

where $a = 4.5 \text{ \AA}$ is the localization, k_B is Boltzmann’s constant and $T = 300 \text{ K}$ [2].

Figure 14.3 illustrated the VRH fitting model for $\text{Nd}_{0.75}\text{Na}_{0.25-x}\text{K}_x\text{MnO}_3$ ($0 \leq x \leq 0.10$) (%). It can be seen that all the experimental data showed a good fitting line with the suggested model indicate the VRH model is a suitable model to use for.

Fig. 14.3 Plot of $\ln \rho$ versus $T^{-1/4}$ for $\text{Nd}_{0.75}\text{Na}_{0.25-x}\text{K}_x\text{MnO}_3$ ($x = 0-0.10$) samples. The solid line indicates the best fit VRH model

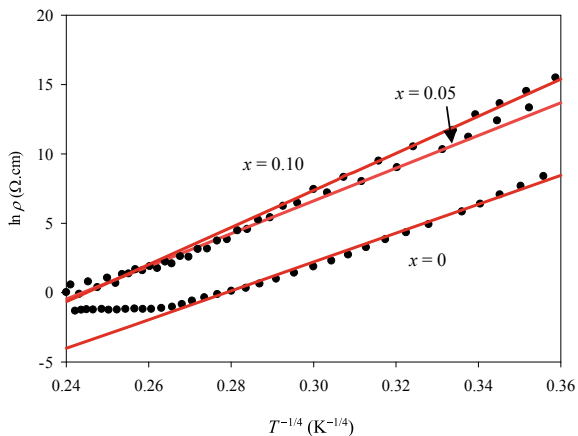


Table 14.1 The square of the linear correlation coefficient (R^2) value for $x = 0$ to 0.10 samples

Sample	R^2	
	VRH	SPH
	$\rho = \rho_{0m} \exp(T_{0m}/T)^{1/4}$	$\rho = BT \exp(E_a/k_B T)$
$x = 0$	99.99	99.30
$x = 0.05$	99.99	99.30
$x = 0.10$	99.99	99.80

Table 14.2 Parameters obtained from VRH model fitting in the insulating region for $\text{Nd}_{0.75}\text{Na}_{0.25-x}\text{K}_x\text{MnO}_3$ ($x = 0-0.10$) samples

Sample	$T_{om}(10^2 \text{ K})$	$N(E_F) (10^{25}) (\text{eV}^{-1} \text{ cm}^{-3})$	$R_h (10^{-8}) (300 \text{ K})$	$E_h (300 \text{ K}) \text{ meV}$
$x = 0$	1.042	2.200	1.295	4.960
$x = 0.05$	1.179	1.940	1.336	5.120
$x = 0.10$	1.337	1.710	1.379	5.280

From Table 14.1, VRH model was designed to explain the resistivity behavior for the insulating region as it produced high R^2 value (99.99%).

Table 14.2 stated all the parameter including Mott's characteristic temperature (T_{om}), density of states at the Fermi level ($N(E_F)$), mean hopping distance (R_h) and hopping energy (E_h) for VRH model.

The stated data in Table 14.2 indicate that the value of internal field $N(E_F)$ reduced from $2.20 \times 10^{25} \text{ eV}^{-1} \text{ cm}^{-3}$ ($x = 0$) to $1.71 \times 10^{25} \text{ eV}^{-1} \text{ cm}^{-3}$ ($x = 0.1$), but then the value of R_h increased from 1.295×10^8 ($x = 0$) to 1.379×10^8 ($x = 0.10$) accompanying with the increasing in E_h from 4.96 meV ($x = 0$) to 5.28 meV ($x = 0.10$) with K^+ substitution. The observation of this behavior could be suggested due to the smaller ionic radius of K^+ ion as compared to Na^+ ion which can be contributed to increasing in localization of charge carriers as a result of lattice distortion in MnO_6 . In fact, similar suggestion were also purposed by Rozilah et al. [2] and Ibrahim et al. [14].

14.4 Conclusion

In conclusion, the influence of K doped in $\text{Nd}_{0.75}\text{Na}_{0.25-x}\text{K}_x\text{MnO}_3$ ($x = 0-0.10$) manganites has been studied. DC electrical resistivity measurements showed all samples exhibited insulating behavior. However, the $d\ln\rho/dT^{-1}$ versus T curve shifted to lower temperature from 210 K to 165 K correspond to weakening of CO state. Further electrical analysis showed that the resistivity behavior in the insulating region for all samples was attributed to VRH where the value of R^2 approaches to 99.99%. The VRH models revealed that the hopping and activation energies decreased as K^+ content increased due to the reduction of MnO_6 octahedral.

Acknowledgements This research was supported by TIER 1 Vot H967 Research Grant from Universiti Tun Hussein Onn Malaysia. Thanks to Ceramic and Amorphous Group, Faculty of Applied Sciences and Technology, Pagoh Higher Education Hub, Universiti Tun Hussein Onn Malaysia, 84600 Panchor, Johor, Malaysia as well as Superconductor Laboratory, Faculty of Applied Sciences, Universiti Teknologi Mara Shah Alam for facilities provided.

References

1. N. Asmira, N. Ibrahim, Z. Mohamed, A.K. Yahya, *Physica B* **544**, 34–46 (2018)
2. R. Rozilah, N. Ibrahim, Z. Mohamed, A.K. Yahya, N.A. Khan, M.N. Khan, *Physica B* **521**, 281–294 (2017)
3. S. Shamsuddin, S.N. Supardan, A.B.M. Ibrahim, A.K. Yahya, *J. Supercond. Novel Magn.* **27**, 1229–1234 (2014)
4. T.L. Phan, N.T. Dang, T.A. Ho, J.S. Rhyee, W.H. Shon, K. Tarigan, T.V. Manh, *J. Magn. Magn. Mater.* **443**, 233–238 (2017)
5. X. Zhang, Z. Li, *J. Rare Earths* **29**, 230–234 (2011)
6. N.N. Ab Mannan, S.A. Razali, S. Shamsuddin, M.Z. Noh, *J. Sci. Technol* **9**, 65–69 (2017)
7. N.N. Khairulzaman, S. Shamsuddin, N. Ibrahim, *J. Adv. Res. Dyn. Control. Syst.* **12**, 759–765 (2020)
8. D. Louca, T. Egami, E.L. Brosha, H. Röder, A.R. Bishop, *Phys. Rev. B* **56**, R8475 (1997)
9. C.H. Chen, S.W. Cheong, *Phys. Rev. Lett.* **76**, 4042 (1996)
10. R.A. Zawawi, N.N. Khairulzaman, S. Shamsuddin, N. Ibrahim, *Int. J. Eng. Technol.* **7**, 75–79 (2018)
11. X. Zhang, L.I. Zhiqing, *J. Rare Earths* **29**, 230–234 (2011)
12. G. Akça, A.O. Ayaş, S.K. Çetin, M. Akyol, A. Ekicibil, *J. Supercond. Novel Magn.* **30**, 1515–1525 (2017)
13. R. Thaljaoui, W. Boujelben, M. Pekała, K. Pekała, J.F. Fagnard, P. Vanderbemden, A. Cheikhrouhou, *J. Magn. Magn. Mater.* **352**, 6–12 (2014)
14. N. Ibrahim, A.K. Yahya, S.S. Rajput, S. Keshri, M.K. Talari, *J. Magn. Magn. Mater* **323**, 2179–2185 (2011)
15. R.A. Zawawi, N.N. Khairulzaman, S. Shamsuddin, *J. Sci. Technol.* **9**, 45–48 (2017)
16. J. Fan, L. Pi, K. He, W. Wang, Y. Zhang, *Europhys. Lett.* **74**, 506 (2006)
17. A. Tozri, J. Khelifi, H. Baaziz, E. Dhahri, E.K. Hlil, *Mater. Lett.* **131**, 61–63 (2014)
18. B. Munirathinam, M. Krishnaiah, U. Devarajan, S.E. Muthu, S. Arumugam, *J. Phys. Chem. Solids* **73**, 925–930 (2012)
19. M.A. Ghani, Z. Mohamed, A.K. Yahya, *J. Supercond. Novel Magn.* **25**, 2395–2402 (2012)
20. S. Mollah, H.L. Huang, H.D. Yang, P. Sudipta, S. Taran, B.K. Chaudhuri, *J. Magnet. Magnet. Mater.* **284**, 383–394 (2004)
21. G. Venkataiah, P.V. Reddy, *J. Magn. Magn. Mater.* **285**, 343–352 (2005)

Chapter 15

Crystalline Structure, Surface Morphology of Perovskite Manganites

$\text{Nd}_{0.75}\text{Na}_{0.25}\text{Mn}_{1-y}\text{Fe}_y\text{O}_3$



S. Shamsuddin, M. Z. H. Mayzan, Fahmiruddin Esa, and L. K. Hao

Abstract The perovskite manganite oxide $\text{Nd}_{0.75}\text{Na}_{0.25}\text{Mn}_{1-y}\text{Fe}_y\text{O}_3$ with $y = 0, 0.10$ and 0.15 samples have been systematically investigated to elucidate the effect of Fe-doped on the crystalline structure and surface morphology. However, study on reporting the relationship between crystalline structure and surface morphology for perovskite manganites is still limited. All samples were prepared using solid-state reaction method and characterized by X-ray Diffraction (XRD) and Scanning Electron Microscope (SEM) with Energy Dispersive X-ray analysis (EDX) to determine the crystal structure and surface morphology as well as identifying the elemental composition. All samples have been observed as an orthorhombic structure with *Pnma* space group and the lattice parameter as well as the unit cell volume had changed due to the Fe doping. The values of unit cell volume were increase consistently from 227.0 \AA^3 ($y = 0$) to 228.1 \AA^3 ($y = 0.15$) which can be suggest that due to the size mismatch between the Fe^{3+} ions and Mn^{3+} ions. The morphological study depicts that the samples have compact and connected grain with well-defined grain boundary and there is not noticeable size change have been shown for all samples. However, the grain size shows an unsystematic trend when the concentration of Fe increase. Apart from that, the EDX measurement shows that all samples have an accurate composition without detected any impurities peak.

S. Shamsuddin (✉) · M. Z. H. Mayzan · F. Esa · L. K. Hao
Ceramic and Amorphous Group, Faculty of Applied Sciences and Technology, Pagoh Higher Education Hub, Universiti Tun Hussein Onn Malaysia, 84600 Panchor, Johor, Malaysia
e-mail: suhadir@uthm.edu.my

M. Z. H. Mayzan
e-mail: zulhilmi@uthm.edu.my

F. Esa
e-mail: fahmir@uthm.edu.my

15.1 Introduction

A lot of researchers had put their attention on investigating the perovskite manganite oxide as they exhibit various interesting and unique properties such as Colossal magnetoresistance (CMR), Jahn-teller effect and charge ordering (CO) [1–3]. According to previous study, the compound of $\text{Nd}_{0.75}\text{Na}_{0.25}\text{MnO}_3$ has orthorhombic structure with $Pnma$ space group [4–6]. In addition, $\text{Nd}_{0.75}\text{Na}_{0.25}\text{MnO}_3$ manganite is a charge ordered manganite and the transition of charge ordering occurred at 170 K (T_{co}) [5, 7]. On the other hand, Fe^{3+} ions have almost identical ionic radius with Mn^{3+} ions which is 0.645 Å [8, 9]. Therefore, the doping of Fe^{3+} ions will not significantly affect the crystal structure as well as the physical properties of the perovskite manganite. The doping of transition element in either A-site or B-site of the perovskite manganite will influence its structural as well as physical properties due to the dopants have different ionic radius with the constituent elements in the perovskite manganite [10–12]. However, the effect of Fe-doped in the Nd-based manganite on the structural along with the physical properties is still limited in the literature. In addition, the density as well as the porosity of the perovskite manganite oxide were also affected by the elemental doping [13]. Thus, a lot of study need to be carried out to understand and reveal the information about the structural and the physical properties of manganites.

Based on the previous study, the surface morphology and microstructure of the perovskite manganite oxide such as grain size displayed a observable change when the dopant is introduced [14, 15]. Previously, the crystalline structure and surface morphology depict similar change upon the elemental doping which suggest that there some connection between the crystalline structure and surface morphology. However, this study is not extensively reported by the researcher. Therefore, this research the Fe^{3+} ions were doped on the $\text{Nd}_{0.75}\text{Na}_{0.25}\text{Mn}_{1-y}\text{Fe}_y\text{O}_3$ to investigate the crystal structure and surface morphology of the samples. Results of density and porosity were also presented and discussed.

15.2 Experiment Methods

The $\text{Nd}_{0.75}\text{Na}_{0.25}\text{Mn}_{1-y}\text{Fe}_y\text{O}_3$ with $y = 0, 0.10$ and 0.15 samples were prepared using solid state reaction method. Stoichiometric amount of Nd_2O_3 , Na_2CO_3 , MnO_2 and Fe_2O_3 oxide powders with high purity of about 99.99% were mix and ground using agate mortar and pestle. After that, the powder was calcinated in air at 1000 °C for 24 h with several intermediate grinding. The powder was then pressed into pellet form with the pressure of 5 tons in 3 min and lastly sintered in air at 1200 °C for 24 h. All samples were characterized by powder X-ray Diffraction (XRD) using Bruker D2 phaser with scanning rate of 2°/min in the range of $20^\circ \leq 2\theta \leq 80^\circ$ and $\text{Cu } \alpha$ radiation at room temperature. The surface morphology and the elemental composition of all samples were determined by Scanning Electron Microscope with

Energy Dispersive X-ray analysis (SEM-EDX) using COXEM EM-30AX PLUS. Bulk density was determined by employing Archimedes principle using acetone as buoyant medium. In this case, the weight of samples in air as well as in acetone were measured using compact precision balance and the bulk density of all samples were calculated using Eq. 15.1

$$\rho_{\text{bulk}} = \frac{w_{\text{air}}}{w_{\text{air}} w_{\text{acetone}}} \rho_{\text{acetone}} \quad (15.1)$$

where ρ_{acetone} is the density of the acetone, w_{air} is the weight of the sample in air and the w_{acetone} is the weight of the sample in acetone. The porosity of all samples was calculated from the standard formula that include bulk density and theoretical density showed in Eq. 15.2

$$\text{Porosity (\%)} = 1 - \frac{\rho_{\text{bulk}}}{\rho_{\text{X-ray}}} 100\% \quad (15.2)$$

where ρ_{bulk} is the bulk density of the sample calculated using Eq. 15.1 and $\rho_{\text{theoretical}}$ is the theoretical density calculated based on the XRD measurement using Eq. 15.3

$$\rho_{\text{X-ray}} = \frac{n'(\sum A_c + \sum A_A)}{V_c N_A} \quad (15.3)$$

where n' is the number of formula unit, $\sum A_c$ is the sum of atomic weight of all cations, $\sum A_A$ is the sum of atomic weight of all anions, V_c is the volume of the unit cell and N_A is the Avogadro's number.

15.3 Results and Discussion

Figure 15.1 showed the powder XRD pattern for the samples $\text{Nd}_{0.75}\text{Na}_{0.25}\text{Mn}_{1-y}\text{Fe}_y\text{O}_3$ with $y = 0, 0.10$ and 0.15 . The analysis of XRD indicates that all samples were indexed in the orthorhombic structure with space group $Pnma$ consistent with previous study by [5, 6]. The XRD results did not depict any discernible impurity peaks as well as secondary peaks suggested that a successful Fe-doped in the compound due the different ionic radius between Fe^{3+} and Mn^{3+} ions. It can be seen from Table 15.1, the lattice parameter for the sample with $y = 0$ were 5.434 \AA , 7.674 \AA and 5.444 \AA for a -lattice, b -lattice and c -lattice respectively while 5.439 \AA (a -lattice), 7.685 \AA (b -lattice) and 5.439 \AA (c -lattice) for $y = 0.10$. For the sample with $y = 0.15$, the a -lattice, b -lattice and c -lattice were 5.443 \AA , 7.685 \AA and 5.452 \AA respectively. Moreover, the unit cell volume of the samples was increase consistently from 227.0 \AA^3 , 227.3 \AA^3 and 228.1 \AA^3 for $y = 0, 0.10$ and 0.15 respectively.

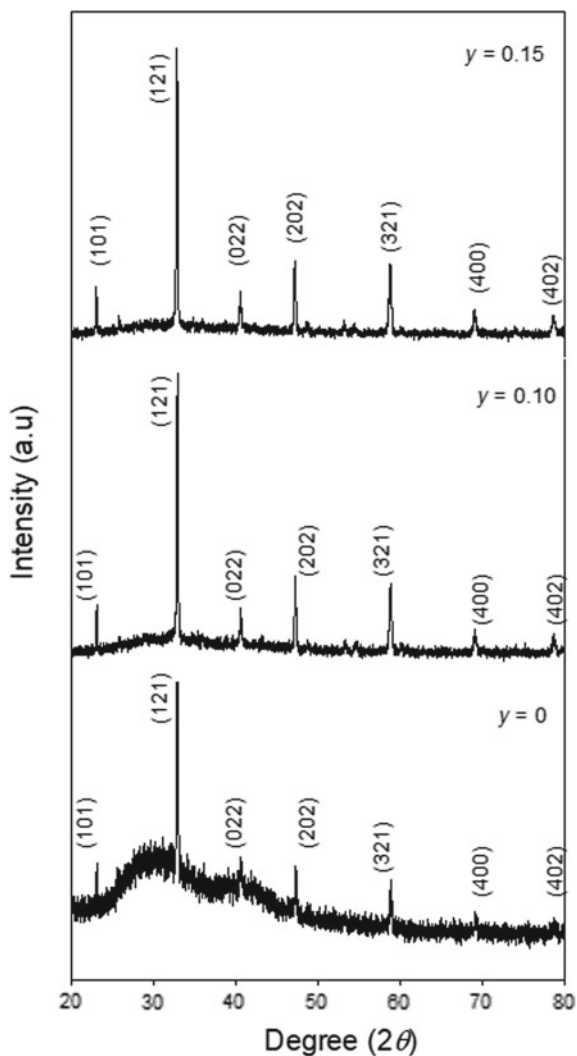


Fig. 15.1 Powder XRD pattern for $\text{Nd}_{0.75}\text{Na}_{0.25}\text{Mn}_{1-y}\text{Fe}_y\text{O}_3$ ($0 \leq y \leq 0.15$)

Table 15.1 Lattice parameters, unit cell volume (V), bulk density (ρ_{bulk}) and porosity of the sample $\text{Nd}_{0.75}\text{Na}_{0.25}\text{Mn}_{1-y}\text{Fe}_y\text{O}_3$ with $y = 0, 0.10$ and 0.15

Sample	Lattice parameter, \AA (± 0.001)			V , \AA^3 (± 0.1)	ρ_{bulk} , g/cm^3 (± 0.001)	Porosity, % (± 0.01)
	a (\AA)	b (\AA)	c (\AA)			
$y = 0$	5.434	7.674	5.444	227.0	5.968	14.20
$y = 0.10$	5.439	7.685	5.439	227.3	5.821	32.15
$y = 0.15$	5.443	7.685	5.452	228.1	5.769	32.53

There is not observable change in crystalline structure for all samples due to the Fe-doped as they have a similar miller indices showed in the XRD pattern. However, there was a slightly change in the lattice parameter and unit cell volume where the *a*-lattice and *b*-lattice displayed an increasing trend while *c*-lattice showed an unsystematic trend upon the Fe doping. The unit cell volume was increasing when the amount of Fe increased. This can suggest that the changes are due to the size mismatch between the Fe^{3+} ions and Mn^{3+} ions [16, 17]. On the other hand, the bulk density showed decreasing trend while the porosity showed increasing trend upon the augmentation of Fe-doping. In this case, the bulk density decreased from 5.970 g/cm^3 ($y = 0$) to 5.769 g/cm^3 ($y = 0.15$) whereas the porosity increased from 14.20% ($y = 0$) to 32.53% ($y = 0.15$). This change can be suggested due to the Fe has larger atomic weight (55.85 g/mol) compared to Mn (54.94 g/mol) [18]. All the lattice parameters, unit cell volume, bulk density, and porosity of the $\text{Nd}_{0.75}\text{Na}_{0.25}\text{Mn}_{1-y}\text{Fe}_y\text{O}_3$ with $y = 0, 0.10, 0.15$ were presented in Table 15.1.

Figure 15.2 showed the SEM images for all samples with the magnification of 5kX. The SEM images revealed that all samples have compact and connected grain with well-defined grain boundary and not observable size change were detected. This can suggest that Fe-doped in the Mn-site will either not change or change fairly the size due to the similar ionic radius of Fe^{3+} ions and Mn^{3+} ions [9]. In addition, the grain size had slightly changed upon the augmentation of Fe doping due to the

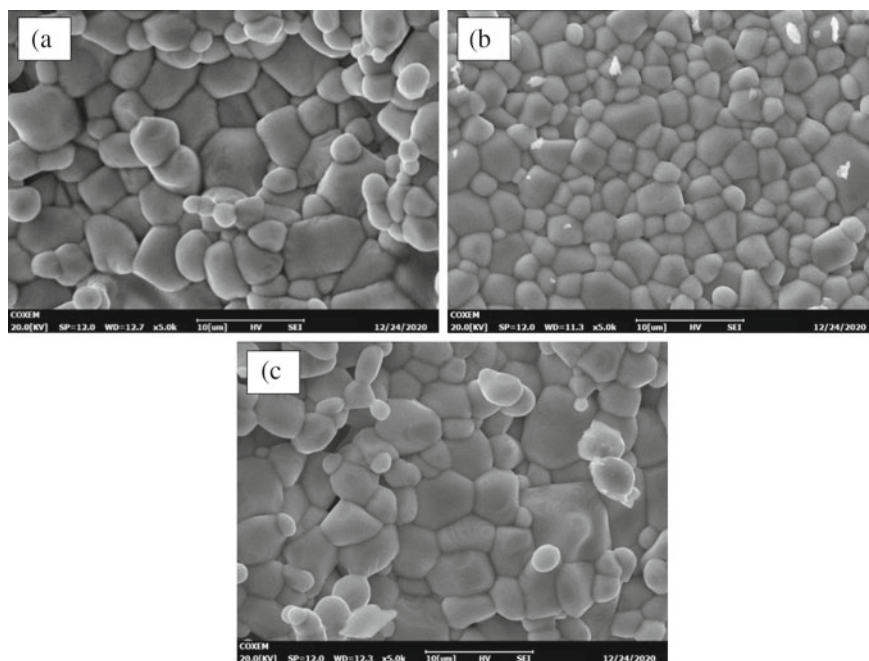


Fig. 15.2 SEM images with 5kX magnification for the sample $\text{Nd}_{0.75}\text{Na}_{0.25}\text{Mn}_{1-y}\text{Fe}_y\text{O}_3$ (a) $y = 0$, (b) $y = 0.10$ and (c) $y = 0.15$

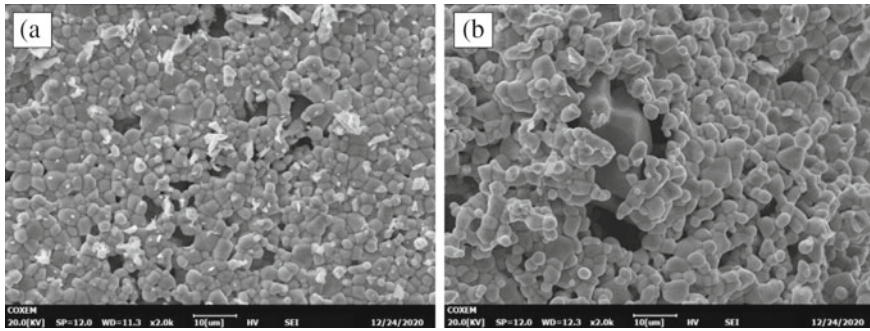


Fig. 15.3 SEM images with 2kX magnification for the sample $\text{Nd}_{0.75}\text{Na}_{0.25}\text{Mn}_{1-y}\text{Fe}_y\text{O}_3$ (a) $y = 0.10$ and (b) $y = 0.15$

size mismatch between Fe and Mn as well as the inhomogeneity of Fe doping in Mn-site [19, 20]. Figure 15.3 also showed the SEM images for all samples with the magnification of 2kX. From the images, the porosity of the sample had slightly increased from $y = 0.10$ to $y = 0.15$ which can be suggested due to the difference in ionic radius between Fe^{3+} and Mn^{3+} ions. In addition, the increasing porosity observed in the image was in line with the calculated porosity.

Figure 15.4 showed the EDX spectrum for all samples. Based on the spectrum, all the constitute elements (Na, Nd, Mn, O, Fe) was detected without existing any impurity peaks indicate that the sample was properly prepared. The detection of the Fe element in the sample of $y = 0.10$ and $y = 0.15$ showed that the Fe was successfully doped into the sample. In addition, the atomic percentage for the Fe was increased from 4.08% to 5.57% for $y = 0.10$ to $y = 0.15$ respectively which was in line with the increasing Fe-doped in the Mn-site of the sample.

According to XRD result, there is not obvious observable change was detected in the crystalline structure of the samples due to the Fe-doped as all samples showed similar XRD pattern with the same miller indices. In addition, the lattice parameter along with the unit cell volume displayed a slightly change upon the Fe-doping. Meanwhile, SEM images did not show any detectable size change when the Fe^{3+} ions doped in the samples. Based on Table 15.1, the porosity had increased when the concentration of Fe increased which was congruent with the result of SEM where the increasing of pores was noticed in the SEM images. Hence, the correlation between the crystalline structure and surface morphology was noticed as X-ray diffraction measurement as well as the Scanning electron microscope study depicted the similar result where there is not apparent change noticed due to Fe-doped.

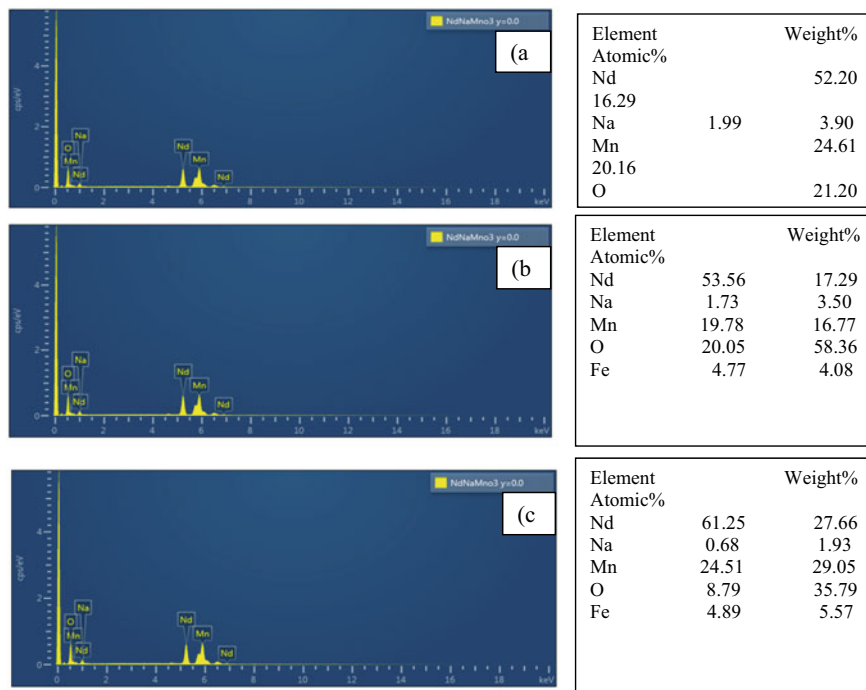


Fig. 15.4 EDX spectrum for the sample $\text{Nd}_{0.75}\text{Na}_{0.25}\text{Mn}_{1-y}\text{Fe}_y\text{O}_3$ (a) $y = 0$, (b) $y = 0.10$ and (c) $y = 0.15$

15.4 Conclusion

In conclusion, XRD measurement on $\text{Nd}_{0.75}\text{Na}_{0.25}\text{Mn}_{1-y}\text{Fe}_y\text{O}_3$ revealed that all samples have orthorhombic structure with $Pnma$ space group and not observable change detected in the crystalline structure of samples which indicate the similar ionic radius between Fe^{3+} and Mn^{3+} ions. The lattice parameter and unit cell volume had change upon the Fe-doping which can be suggested due to the size mismatch between Fe^{3+} and Mn^{3+} ions. The SEM study showed that all samples had compact and connected grains with well-defined boundary. There is not observed size change for all samples which indicate the similar ionic radius of Fe^{3+} and Mn^{3+} ions. EDX analysis indicate that all samples have exact composition without any impurity peak. The density of the samples had decreased due to Fe-doped which suggest that Fe has larger atomic weight that Mn. The porosity displayed increasing trend which is line with the porosity noticed in the SEM image. Hence, the correlation between crystalline structure and surface morphology had been revealed based on the results obtained and indicate the effect of Fe-doped in the samples.

Acknowledgements This research was supported by TIER 1 Vot H967 Research Grant from Universiti Tun Hussein Onn Malaysia. The author would like to thank the Ceramic and Amorphous

Group, Faculty of Applied Sciences and Technology, Pagoh Higher Education Hub, Universiti Tun Hussein Onn Malaysia for the facilities provided.

References

1. S Shamsuddin ABMA Ibrahim AK Yahya 2013 *Ceram. Int.* 39 185 188
2. NN Khairulzaman S Shamsuddin N Ibrahim 2020 *J. Adv. Res. Dyn. Control Syst.* 12 759 765
3. S Heini M Boudard S Zemni M Oumezzine 2014 *Ceram. Int.* 40 16041 16050
4. SA Razali N Ibrahim S Shamsuddin MZ Noh 2018 *Int. J. Eng. Technol.* 7 85 88
5. RA Zawawi N Khairulzaman S Shamsuddin 2017 *J. Sci. Technol.* 9 45 48
6. N.N. Ab Mannan, S.A. Razali, S. Shamsuddin, M.Z. Noh, *J. Sci. Technol.* **9**, 65–69 (2017)
7. X Zhang Z Li 2011 *J. Rare Earths* 29 230 234
8. NV Khiema LV Bau NM An NX Phuc DNH Nam 2003 *Physica B* 327 187 189
9. Y. Zhou, X. Zhu, S. Li, *Ceram. Int.* **43**, 3679–3687 (2017)
10. W. Boujelben, M. Ellouze, A. Cheikh-Rouhou, R. Madar, H. Fuess, *Phys. Status Solidi (A) Appl. Res.* **201**, 1410–1415 (2004)
11. W. Cheikh-Rouhou Koubaa, M. Koubaa, A. Cheikh-Rouhou, *J. Magnet. Magnet. Mater.* **316**, 648–651 (2007)
12. P Thamilmaran M Arunachalam S Sankarajan K Sakthipandi 2015 *J. Magn. Magn. Mater.* 396 181 189
13. REA Ngida MF Zawrah RM Khattab E Heikal 2019 *Ceram. Int.* 45 4894 4901
14. R.A. Zawawi, N.N. Khairulzaman, S. Shamsuddin, N. Ibrahim, *Int. J. Eng. Technol.* **7**, 75–79 (2018)
15. S Kossi El S Ghodhbane S Mnefgui J Dhahri EK Hlil 2015 *J. Magn. Magn. Mater.* 395 134 142
16. S Singh D Singh 2017 *J. Alloy. Compd.* 702 249 257
17. J Dhahri A Dhahri M Oumezzine EK Hlil 2015 *J. Magn. Magn. Mater.* 378 353 357
18. AAS Hassan W Khan S Husain P Dhiman M Singh 2020 *Int. J. Appl. Ceram. Technol.* 17 2430 2438
19. I.A. Abdel Latif, *J. Phys.* **1**, 15–31 (2012)
20. I Hussain MS Anwar SN Khan E Kim BH Koo 2017 *J. Nanosci. Nanotechnol.* 17 7859 7864

Chapter 16

Element Identification of Different Gemstones by Using LIBS-PCA Method



Noradilah Mohamad Saufi, Nurul Nadia Adnan, Syed Zuhaib Haider Rizvi, Mohd Sufi Roslan, Siti Norazida Razak, Nurhafizah Hasim, and Nurin Iffah Mansor

Abstract Laser-induced breakdown spectroscopy (LIBS) is a flexible non-destructive method for both qualitative and quantitative analysis. These days, people are obsessed with gemstones apart from gold caused by the attractive natural look and, therefore, the sort of colour that is limited availability for the gemstone. However, due to this demand, the price was not stable. The seller easily misled the consumer because it is harder to distinguish the quality of different original natural gemstones. This study aims to characterize and analyze the element that is present from each spectrum in the gemstone. After that, the main element of the gemstone was identified, and the gemstone was classified by using multivariate analysis, Principal Component Analysis (PCA). Four types of gemstones were employed in this study. The spectrum was produced, and each line of emission represented the elements contained in the gemstone. LIBS-PCA technique is a method that can be used for the classification and characterization of gemstones.

N. M. Saufi · N. N. Adnan (✉) · S. Z. H. Rizvi · N. I. Mansor
Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, Pagoh Educational Hub, 84600 Pagoh, Johor, Malaysia
e-mail: nadia@uthm.edu.my

N. N. Adnan · S. Z. H. Rizvi
Photonics Devices and Sensor Research (PDSR), Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, Pagoh Educational Hub, 84600 Pagoh, Johor, Malaysia

M. S. Roslan · S. N. Razak
Department of Science and Mathematics, Center for Diploma Studies, Universiti Tun Hussein Onn Malaysia Educational Hub Pagoh, 84600 Muar, Johor, Malaysia

S. Z. H. Rizvi · M. S. Roslan
Microelectronic and Nanotechnology Shamsuddin Research Center (MiNT-SRC), Universiti Tun Hussein Onn Malaysia, 86400 Parit Raja, Batu Pahat, Johor, Malaysia

N. Hasim
Advanced Optical Material Research Group (AOMRG), Department of Physics, Faculty of Science, Universiti Teknologi Malaysia, 81310 Bahru, Johor, Malaysia

16.1 Introduction

Nowadays, people were obsessed with gemstones apart from gold caused by the attractive natural looked and therefore the sort of colour that was limited availability for the gemstone. A natural or real gemstone such as a mineral, stone or organic matter which will be cut and polished but otherwise treated to be used for jewellery or the other accessory. The precious gemstone had durability, beauty and rarity, whereas the semiprecious gemstones only have one or two qualities. However, the thing that they had connected was that the price. The harder the possibility to had the colour, the upper the cost of the gemstone. Topaz, Corundum (Ruby and Sapphire), Beryl (Emerald and Aquamarine), Diamond and Opal were commonly classified as precious stones. In contrast, all other gemstones were classed as a semiprecious stones.

A gemstone was a naturally occurring crystalline form in shape, and different colours of gemstone had different types of chemical elements or impurities that were present. The gemstone could be found inside the earth's crust gradually from time to time in a few states of the world. These places combined with historically important gem deposits had not always provided the best quality material in enough amounts to achieve the current demand for gemstones among consumers. Usually, little quantity of high-quality gem material typically formed at a given mining locality and recovered a much more percentage of lowered quality material had little market value. As a solution, these individuals will continue to strive to built techniques to treat this lower-quality material in the laboratory to increase its features and marketability for gem purposes.

In the last few decades, gemstone testing had grown from rather basic characterization of physical properties (for example, density, refraction indices, dispersion, absorption, hardness and fluorescence) into advanced material science using a refined scientific technique for detailed chemical structural analysis. For instance, LA-ICPMS could distinguish almost all chemical elements with a limit of detection with a range of parts per million (ppm) to even parts per billion (ppb) levels [1, 2]. Besides that, LA-ICP-TOF-MS was a sensitive technique for the chemical analysis of gemstones and pearls compared to the LA-ICPMS, which yields high-quality elemental analyses, results in multidimensional data, and feeds the database for statistical analysis [3]. Other than that, characterizing pearls structure by using X-ray phase-contrast and neutron imaging [4].

Laser-induced breakdown spectroscopy (LIBS) was a spectroscopic technique for analyzing the material's elements by utilizing quantitative and qualitative methods [5]. This approach examined plasma emission generated by focusing high power laser pulses on the samples with atmospheric pressure conditions. LIBS was also an atomic emission spectroscopy method that employed high energy laser pulses to incite optical sample excitation [6]. The relation between samples and focused laser pulses results in the plasma was composed with ionized matter [7]. The electrons in the excited atoms will be backed to their normal state or ground state by emitting photons with a specific wavelength. The photon emission observed by the function of

the total intensity corresponds to the concentration or the number of atoms (elements) that exist in the sample. The quality of analytical results was highly dependent on the plasma formation and detection process [8].

Analytical methods of atomic spectroscopy have been used in a bigger range of scientific applications. Nowadays, Laser-Induced Breakdown Spectroscopy (LIBS) has become a well-known analytical method used to characterize and identify materials. It is multi-elemental analysis, fast, remote sensing, no need for sample preparation, low cost and ease of use [9]. PCA is firmly related to factor analysis. Factor analysis includes additional domain-specific assumptions about the underlying structure and solves the eigenvectors of a moderately different matrix. Hotelling [10] change in the quality control variable also changes factor analysis about the contrast between factor analysis and PCA of Jolliffe's Principal Component Analysis [11]. It can be done by decomposition of eigenvalue or single value of a data correlation matrix after the first data set undergo normalization [12]. Results also discussed that in terms of loading or the weight to get the component score, each original standardized variable should be multiplication [13]. In this study, element identification of different gemstones was carried out by combining laser-induced breakdown spectroscopy (LIBS) technique with multivariate analysis tools and principal component analysis (PCA).

16.2 Materials and Method

Figure 16.1 shows a schematic illustration of the experimental setup used in this research. Q-switched Nd: YAG laser with a wavelength of 1064 nm, pulse width and frequency of 1 Hz/s was employed in this study. A plano-convex lens (L1), convex

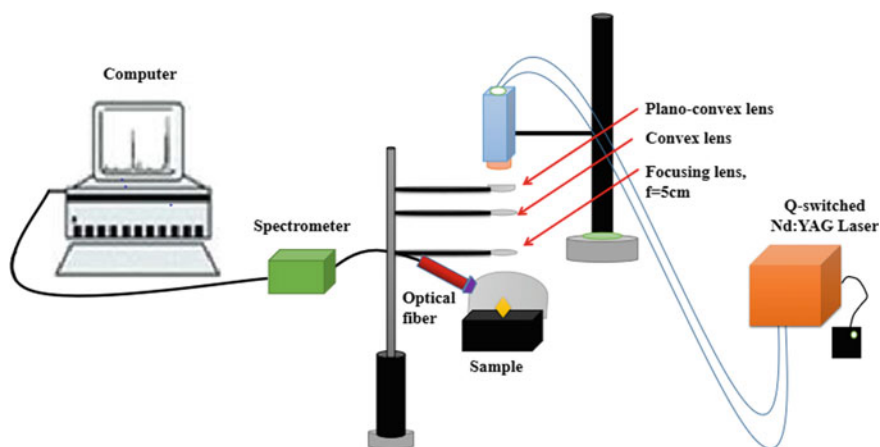


Fig. 16.1 LIBS schematic experiment setup

lens (L2) and focusing lens (L3) with focal lengths 5 cm, 9 cm and 5 cm, respectively, were used to form the best focus point of the laser. In this study, the samples used are Tourmaline, Sapphire, Topaz, Emerald and Amethyst. High energy laser (300 mJ until 1 kJ) is targetted on the gemstone surface through L1, L2 and L3 lenses. Then, it causes the sample's breakdown and the plasma's formation. At angle 45° , fibre optic captured the plasma and sent it to the spectrometer to convert into a spectrum graph in the origin software. The computation was carried out with five shots for each sample. Then, the average value of these measurements will be used as the data set in the data analysis. Each peak intensity in the spectrum was matched with the National Institute of Standards and Technology (NIST) database to identify the elemental content found in the sample based on the samples' expected molecular element.

16.3 Results and Discussion

LIBS-PCA method is the best method to detect and analyze the type of element present in the gemstones. The line of spectrum form shows the different existence of an element. Principle Component Analysis (PCA) is a method considered a classical characteristics extraction and data visualization method. This technique can be used to highlight the variation and strong patterns of data set. Figure 16.2 shows the emission spectrum from gemstones.

Table 16.1 shows the element's content on gemstones. The element with wavelength range 600–900 nm was identified by referred to NIST database. Wavelength and element in tourmaline are OI at 672.68 nm and OII at 688.74 nm, 841.40 nm, and 881.88 nm. Element Silicon (Si), Oxygen (O), Aluminium (Al), with trace element Fluorine (F) and Hydrogen (H) were the expected content to be found in this topaz sample. From the NIST database, two wavelengths can be identified from 650 to

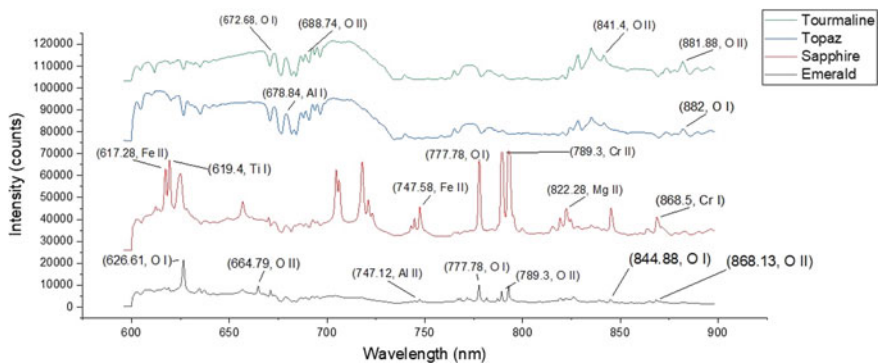


Fig. 16.2. LIBS spectra for four gemstones samples were acquired using optical emission spectroscopy with elemental mapping from the NIST database

Table 16.1. Element contained in the gemstone sample

Gemstone	Element	Peak (nm)
Tourmaline	O I	672.68
		688.74
		841.40
		881.88
Topaz	Al I	678.84
	O I	882.00
Sapphire	Fe II	617.28
		747.58
	Ti I	619.40
	O I	777.78
	Cr II	789.30
	Mg II	822.28
	Cr I	868.50
Emerald	O I	626.61
		777.78
		844.88
	O II	664.79
		789.30
		868.13
	Al II	747.12

900 nm: AlI at 678.84 nm and OI at 882 nm. The expected element content in the Sapphire sample is Oxygen (O), Aluminium (Al) with trace amounts (α) of elements. For instance, Iron (Fe), Titanium (Ti), Chromium (Cr), Vanadium (V) and Magnesium (Mg). The wavelength of FeII 617.28 nm, TiI 619.4 nm, FeII 747.58 nm, OI 777.78 nm, CrII 789.3 nm, MgII 822.28 nm, and CrI 868.55 nm represent the elements contained in the sapphire sample that are matched with the NIST database. In the emerald sample, the expected element's content is Silicon (Si), Oxygen (O), Beryllium (Be) and Aluminium (Al). Seven wavelengths in this spectrum represent the elements contained in the emerald sample. OI 626.61 nm, OII 664.79 nm, AlII 747.12 nm, OI 777.78 nm, OII 789.3 nm, OI 844.88 nm and AlII 868.13 nm. LIBS provides a simple method to identify a sample's elements by capturing the emission of a plasma.

16.4 Principle Component Analysis of LIBS Spectra

PCA was a multivariate data analysis to summarize the information contained in large data tables to a smaller set to be more easily visualized and analyzed. PCA helps to identify the correlations between data points, such as whether there is a correlation between different gemstones.

The loading plot shows a strong relationship between each characteristic and principal component. Loading plots indicate the actual dimension of the problems that PCA transforms. The loading plot of PC for all gemstones in this research was shown in Fig. 16.3. The initial principal component explains 72% of the variation, 11% explained by the second.

Figure 16.4 display the scree plots of PCA analysis for gemstones. Scree plot is a line segment plot that shows the eigenvalue for each PC (Principal Component) (PC1,

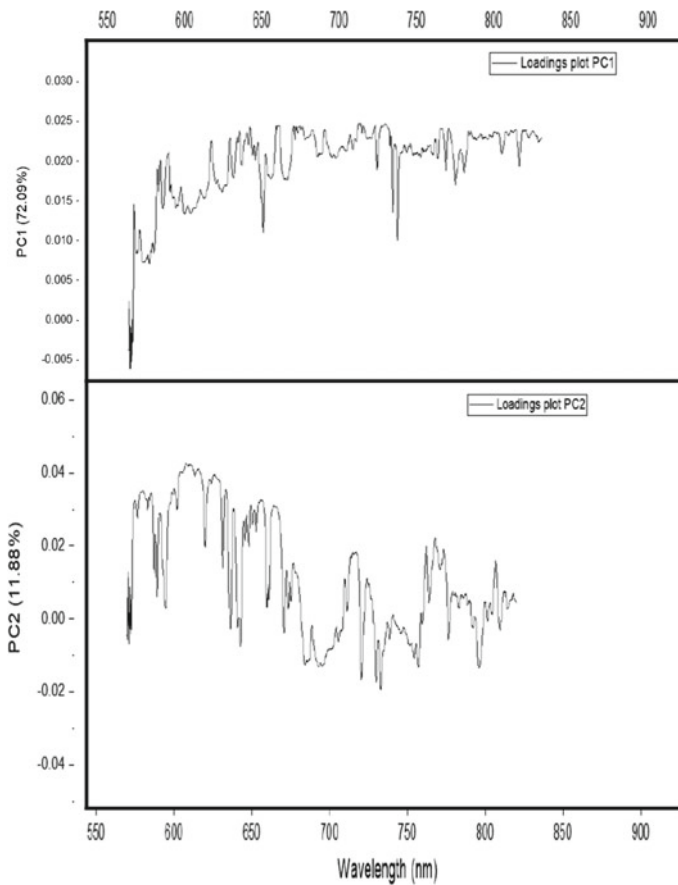


Fig. 16.3 Loading plots for principal components of gemstones

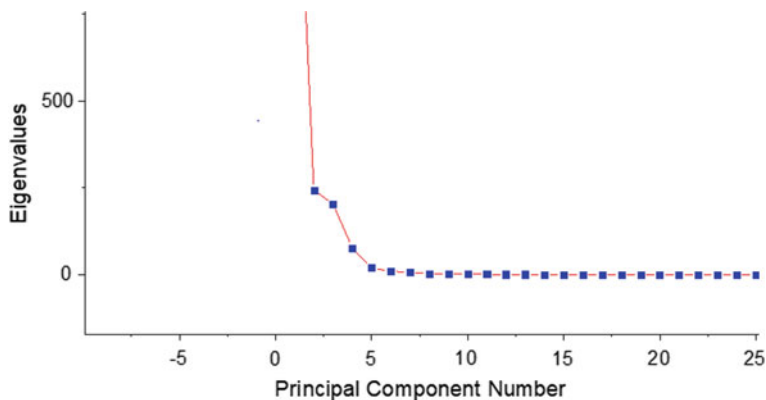


Fig. 16.4 Scree plot for principal component analysis for gemstones

PC2 and PC3). The scree plot portrayed the variety of each principal component that was captured from the data. The scree plot is a steep curve that bends at an “elbow” where the breaking off point and then flattens out. That condition occurs due to sufficient PCs (for first two or three) to represent the essence of data. The number of PCs before this elbow point is considered sufficient to describe the data variance because the later PCs do not cover much variance. In other words, it helps to decide how many data dimensions are needed and exclude them for further analysis.

The score plot was used to demonstrate each sample’s projection that was observed to the principal components. The score plot gemstones are referring to the different types of gemstones.

Figure 16.5 illustrates the location of gemstones in the multivariate region for score vectors of the three principal components. The outcome is clustered into four groups according to the types of gemstones: emerald, sapphire, topaz, and tourmaline. The PCA method had proved the effectiveness to discriminate the different types of gemstones. LIBS-PCA methods offer a rapid identification, classification and differentiation of LIBS spectrum, which has correlated variables toward each other.

16.5 Conclusion

LIBS-PCA shows that this method can be used to identify the gemstones samples’ element fast, easy and at low cost. It can be seen through the visual plasmas and their emission spectra form. From visual plasma, spectrum graphs were produced showing differences in relative intensity. Identification of elements in the spectrum was made using the NIST or National Institute of Standard and Technology database based on the wavelength of each of the peaks that will determine the type of element and its characteristics. Four clusters were observed to differentiate the types of gemstones. LIBS-PCA is a simple method to discriminate the different types of gemstones.

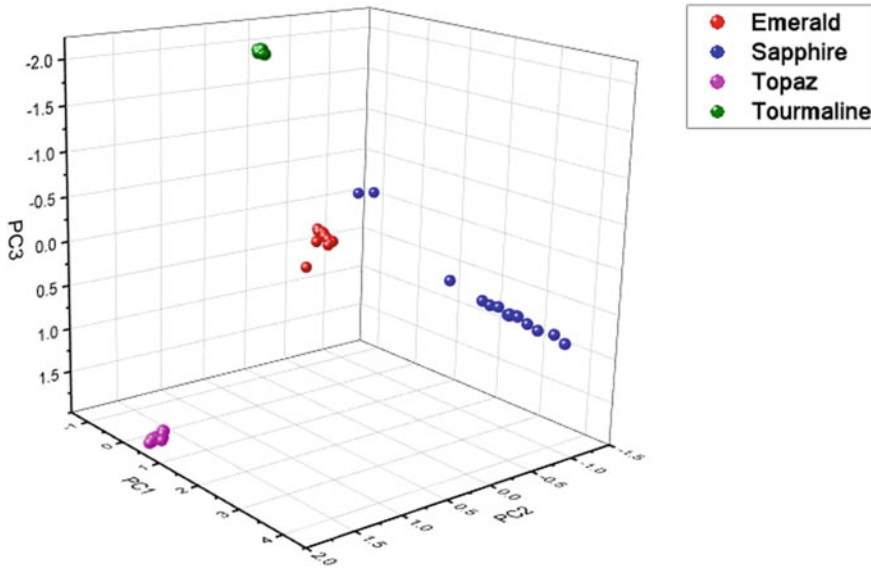


Fig. 16.5 PCA plot of gemstones samples

Acknowledgement This research was supported by the Ministry of Higher Education (MOHE) through Fundamental Research Grant Scheme for Research Acculturation of Early Career Researchers (FRGS-RACER) (RACER/1/2019/STG02/UTHM//1) and Universiti Tun Hussein Onn Malaysia Postgraduate Research Grant (H632) for funding the project. Special thanks from author for Faculty of Applied Sciences and Technology (FAST) of Universiti Tun Hussein Onn Malaysia, give full support in terms of instruments.

References

1. A. Abduriyim, H. Kitawaki, Applications of laser ablation—inductively coupled plasma—mass spectrometry (LA-ICP-MS) to gemology. *Gems Gemology* **42**(2), 98–118 (2006)
2. D. Stefas, N. Gyftokostas, S. Couris, Laser induced breakdown spectroscopy for elemental analysis and discrimination of honey samples. *Spectrochim. Acta Part B At. Spectrosc.* **172**, 105969 (2020)
3. H.A.O. Wang, M.S. Krzemnicki, J.P. Chalain, P. Lefèvre, W. Zhou, L. Cartier, Simultaneous high sensitivity trace-element and isotopic analysis of gemstones using laser ablation inductively coupled plasma time-of-flight mass spectrometry. *J. Gemmol.* **35**(3), 212–223 (2016)
4. D. Micieli, D. Di Martino, M. Musa, L. Gori, A. Kaestner, A. Bravin, A. Mittone, R. Navone, G. Gorini, Characterizing pearls structures using X-ray phase-contrast and neutron imaging: A pilot study. *Sci. Rep.* **8**(1) (2018)
5. A. Khumaeni, W. Setia Budi, K.S. Firdausi, Penghitungan Rasio Intensitas Ca (II) 396,8 nm dan Ca (I) 422,6 nm pada Sampel Tasbih Asli dan Imitasi menggunakan Metode laser induced shock wave plasma (LISPS). *Berkala Fisika* **9**(2), 55–62 (2012)

6. B. Kearton, Y. Mattley, Laser-induced breakdown spectroscopy: Sparking new applications. *Nat. Photonics* **2**(9), 537–540 (2008)
7. D.A. Cremers, L.J. Radziemski, *Handbook of Laser-Induced Breakdown Spectroscopy*, 2nd edn. (John Wiley and Sons, 2013)
8. H. Suyanto, In *Qualitative Analysis of Black Stone and Its Application for Detecting Ag and Pb in Liquid Sample by Laser-Induced Breakdown Spectroscopy (LIBS)*. AIP Conference Proceedings, vol. 1719, no. 1. (2016), p. 030049
9. K.G. Suastika, H. Suyanto, S. Gunarjo, Characterization of Amethysts from Sukamara, Central Kalimantan, using laser-induced breakdown spectroscopy (LIBS). *J. Phys. Soc. Indonesia* **1**(2), 5–8 (2020)
10. H. Hotelling, Analysis of a complex of statistical variables into principal components. *J. Educ. Psychol.* **24**(6), 417–441 (1933)
11. R.R. Sundararajan, Principal component analysis using frequency components of multivariate time series. *Comput. Stat. Data Anal.* **157**, 107164 (2021)
12. H. Abdi, L.J. Williams, Principal component analysis. *Wiley Interdisc. Rev. Comput. Stat* **2**(4), 433–459 (2010)
13. N.J. Cox, Book review: Multivariate statistics for the environmental sciences. *Prog. Phys. Geogr. Earth Environ.* **28**(3), 460–461 (2004)

Chapter 17

Discrimination of Precious and Semi-Precious Gemstones Using Laser-Induced Breakdown Spectroscopy and Machine Learning Approaches



Ain Zamira, Syed Zuhaib Haider Rizvi, Mohd Sufi Roslan, Nurul Nadia Adnan, Siti Noraiza Ab Razak, and Muhammad Arif Jalil

Abstract Laser-Induced Breakdown Spectroscopy (LIBS) is a very simple but capable spectroscopic analytical method that performs simultaneous multi-element analysis in a single laser shot. By using LIBS, we can identify artificial gemstones that look like real ones. In this study, the samples that are used include five different gemstones namely Sapphire, Emerald, Amethyst, Tourmaline, Topaz, and two artificial replicas of Amethyst and Emerald that closely resemble their original counterparts. The LIBS spectra were collected from each of the gemstones by ablating them with a Q-Switched Nd: YAG Laser of wavelength 1064 nm, pulse width (6–10 ns), and maximum energy of 1000 mJ. The radiation from the resulting plasma was recorded with a compact miniature USB2000+ spectrometer in the wavelength range of 600–900 nm. Each sample underwent at least five cleaning laser shots before 10 measurement shots. Two different datasets were prepared namely dataset 1, comprising full-length spectra, and dataset 2, comprising spectral lines only. We

A. Zamira · S. Z. H. Rizvi (✉) · N. N. Adnan
Center of Research (CoR) Photonics Devices and Sensors Research Center (PDSR), Universiti Tun Hussein Onn Malaysia, 84600 Pagoh, Johor, Malaysia

A. Zamira
e-mail: syedzuhaib@uthm.edu.my

S. Z. H. Rizvi · N. N. Adnan
Department of Physics and Chemistry, Faculty of Applied Science and Technology (FAST), Universiti Tun Hussein Onn Malaysia, 84600 Pagoh, Johor, Malaysia

S. Z. H. Rizvi · M. S. Roslan
Microelectronic and Nanotechnology—Shamsuddin Research Center (MiNT-SRC), 86400 Parit Raja, Johor, Malaysia

M. S. Roslan · S. N. A. Razak
Center for Diploma Studies (CeDS), Universiti Tun Hussein Onn Malaysia, 84600 Pagoh, Johor, Malaysia

M. A. Jalil
Physics Department, Faculty of Science, Universiti Teknologi Malaysia, 81310 Skudai, Johor Bahru, Johor, Malaysia

applied Principal Component Analysis (PCA) on LIBS spectral data to visualize its discriminability and Random Forest (RF) to perform classification. Scree plots describe the discriminability of samples generated in PCA and suggested PC1 and PC2. Loading plots showed the most significant spectral regions for PC1, PC2, and PC3 which covered the data variance of 67.95%, 20.16%, and 6.28% respectively. For the RF classification accuracy, the best splitting for training and testing sets was 80–20%. It demonstrated a perfect 100% accuracy in discriminating gemstones with both dataset 1 and dataset 2.

17.1 Introduction

A gemstone is a piece of mineral crystal that is used in jewelry or other decorations and ornamental pieces in cut and polished form. Most gemstones are hard, but due to their sheen or other physical properties that have artistic value, some soft minerals, for example, amethyst or quartz are used in jewelry. The precious stones in common use are diamond, ruby, sapphire, and emerald, and some of the semi-precious gemstones are amethyst, topaz, tourmaline among many others. Gemologists identify gemstones in modern times. A gemologist first uses its chemical composition to classify a gemstone [1]. Diamonds, for example, are made of carbon (C), and rubies of aluminum oxide (Al_2O_3). Gems are categorized into diverse groups, types, and variabilities. As ruby, is a red family of the corundum, while any other color of corundum is known as sapphire [1]. Samples of such variants of beryl include green (green), aquamarine (blue), gold-beryl (red), goshenite (colorless), heliodor (yellow), and morganite (pink) [2].

Differences in chemical compositions of gemstones allow their classification rapidly using laser-induced breakdown spectroscopy (LIBS) which is a rapid analytical technique and does not require sample preparation. It provides many convincing advantages over other analytical techniques. For instance, the capability of analysis without sample preparation; extremely quick measurement times (normally only a few seconds) for a solitary spot analysis; multi-element analysis [3]. Light elements like Li, C, B, and Na in particular can be detected by LIBS but not by XRF [4]. Even so, for light elements such as Mg, Al, and Si, LIBS has better detection limits [5]. LIBS can be utilized for the detection of heavy metal elements also [6]. The standard LIBS detection threshold is small, i.e., ~ppm. For a variety of sample matrices, including metals, semiconductors, lenses, biological tissues, insulators, fibers, soils, thin-painted structures, electronic materials, etc. [3]. The wide range of elements detection with LIBS and the simplicity of the measuring setup offers sufficient motivation to use LIBS for quick scanning purposes.

The repetition rate for a pulsed laser (or pulse repetition frequency) is defined as the number of emitted pulses per second, or more precisely the inverse temporal pulse spacing. In this research, Q-switching of solid-state lasers is normally used, it typically allows repetition rates from below 1 Hz to the order of 100 kHz. For active Q switching, the pulse repetition rate is determined by an external drive signal, while

in passive Q switching it depends on the magnitude of loss modulation, the pump power, and various other parameters. The repetition rate can influence the results when used a higher repetition mode [7].

By using LIBS, we can identify the artificial gemstones that were treated to look like real stones. So, it can determine whether the gemstones are real or synthetic. A real gemstone is a gemstone that is grown or produced in a natural world without the influence of human activity. Fake gemstones are gemstones that are not real and are usually made of glass or a polymer while synthetic gemstones are grown gemstones that look like natural gemstones without being chemically and optically identical [8]. There are other methods or techniques that are commonly used to identify gemstones. For instance, Laser ablation inductively coupled plasma mass spectrometry (LA-ICP-MS) is a solid sample introduction device that uses a laser beam to vaporize and ablate a solid sample and Secondary ion mass spectrometry (SIMS) is a technique used to analyze the composition of solid surfaces and thin films by sputtering the surface of the specimen with a focused primary ion beam and collecting and analyzing ejected secondary ions [4].

The primary aim of this project is to identify various gemstones based on their LIBS spectra and to discriminate between real and fake ones. This will be done by picking up small differences in the LIBS spectra of stones (which arise due to small differences in the chemical composition of the stones) by using statistical methods applied with machine learning techniques, Principal Component Analysis (PCA), and Random Forest (RF).

17.2 Materials and Methods

In this study, the samples that are used include five types of gemstones Sapphire, Emerald, Amethyst, Tourmaline, Topaz, and two synthetic stones which are Fake Amethyst and Fake Emerald. These gemstones are obtained from a local gem trader and got certification from Global Gemological Laboratory in Kuala Lumpur.

Figure 17.1 displays the schematic diagram of LIBS experimental setup for producing laser-induced plasma on gemstones samples and collects the light emissions using an optical spectrometer. LIBS is a well-known spectroscopic technique of a strongly ionized laser-induced plasma. By focusing a high-power laser pulse directed on a substance such as solid, liquid, gas, and aerosols, the focal point produces rapid local heating and intense evaporation followed by the formation of plasma. The interaction between a laser beam and a solid is a complex process that depends on several laser and solid material characteristics. Plasma evolution is depending on laser power, its wavelength, pulse duration, target vapor concentration, concentration of atmospheric gas, and heat.

A Q-Switched Nd: YAG Laser of wavelength 1064 nm, pulse width (6–10 ns), and maximum energy 1000 mJ were used to ablate samples. The laser beam was reflected from a mirror surface and focused on the sample surface by a focusing lens of focal length 100 mm. The focusing lens was cleaned using isopropyl alcohol to

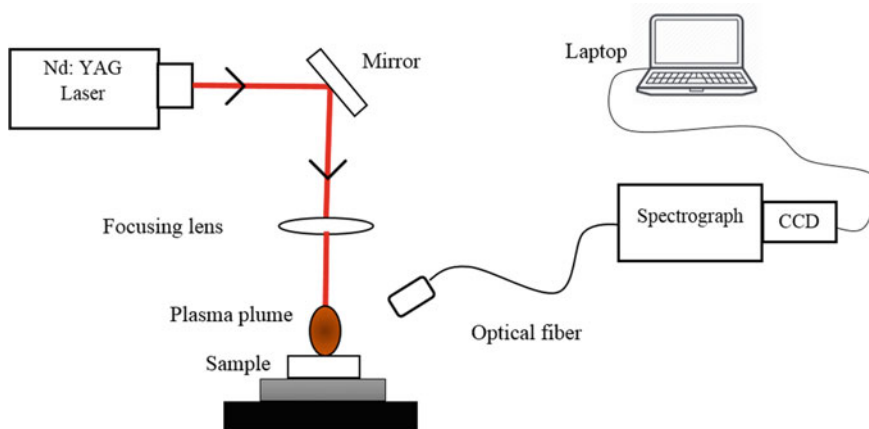


Fig. 17.1 Schematic illustration of experimental set-up

make certain the lens is clean and free from dust particles and skin oils. The radiations emitted from plasma were recorded with USB2000+ (600–900 nm) Charge Coupled Device (CCD) spectrometer. For each gemstone sample, the sample underwent at least five cleaning laser shots before the 10 measurement shots. This is to make certain that the spectral data is obtained from the samples and does not include any interference from surface contaminants. Plasma emission spectrums were saved on the laptop for each laser shot.

The spectral lines observed in the spectra were identified with the help of Atomic Spectral Database of the National Institute of Standards & Technology (ASD-NIST). OriginPro 2018 and Python are used to analyze the spectra and implement classification algorithms respectively.

17.2.1 Classification

In this study, two chemometric methods, Principal Component Analysis (PCA) and Random Forest (RF) are employed for the classification of gemstones.

PCA is used as an exploratory data analysis for showing the multivariate data arrangement. PCA offers useful schemes by describing the first major components (PCs), a linear combination of original variables, preserving the full spectral variability amount [9]. Thus, PCA allows simple visualization of sample distribution, detection of irregular spectra (outlier), and variable deduction. Here the use of PCA offers a quick and easy way to visualize and detect these outliers and therefore PCA filtering outliers is an efficient way of improving the data. PCA has helped to visualize the discriminability of our data and highlighted the spectral regions that carry the most significant discriminatory features.

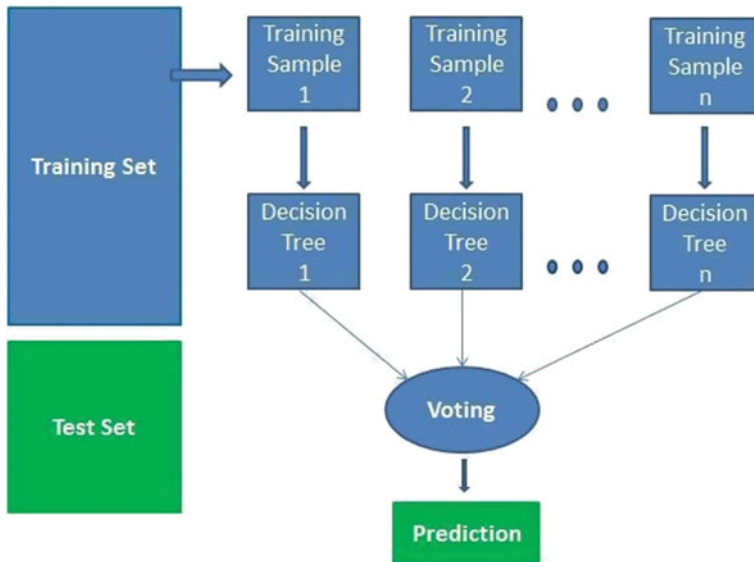


Fig. 17.2 Random forest algorithm [10]

Random Forest (RF) is a classification method based on statistical learning theory. There are two important parameters in RF: (1) the number of trees in the forest (n features) and (2) the dataset that splits each node. Figure 17.2 shows the random forest algorithm which works as a large collection of décor-related decision trees. The random forest Algorithm works in four steps, first, to pick random samples from a specific dataset. Next, for every sample, this algorithm will create a decision tree. Then every decision tree produces a prediction result. At this stage, any predicted result will be voted on. Finally, the algorithm picks the most voted result of the forecast to give out the final prediction.

17.3 Results and Discussion

17.3.1 Data Analysis

In this analysis, laser-induced emission spectra from different gemstones were collected using Nd: YAG laser and USB2000+ (600–900 nm) wavelength range. Each of the gemstone samples was analysed from 5 different spots on the surface of the gemstones and these spectra were averaged to obtain 10 spectra from each spot. Therefore, a total of 50 spectra was recorded for each sample. Figure 17.3, presents the emission spectrum of Amethyst, Fake Amethyst, Emerald 1.67 carats, Fake emerald, Sapphire 1.60 carats, Topaz and Tourmaline 2.95 carats respectively.

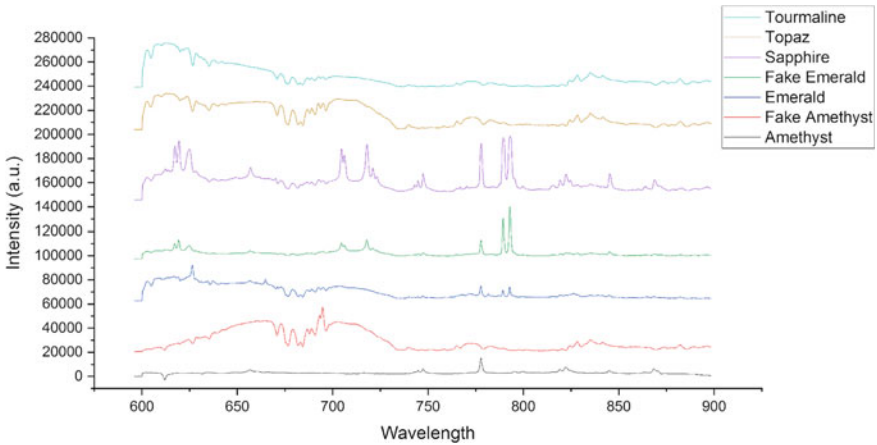


Fig. 17.3 LIBS spectrum of Amethyst, Fake Amethyst, Emerald 1.67 crt, Fake emerald, Sapphire 1.60 crt, Topaz and Tourmaline 2.95 crt

17.3.2 Loadings Plots

Loading plots of PCA shows the correlation of PCs and the spectral features. More raw spectral information is covered by PCs with large loading values, and the corresponding wavelength variable is of greater significance. Positive loading values suggest a direct correlation between PC and the wavelength, while negative loading value refers to a negative correlation between the PC and the wavelength. PCA loading plots are then commonly used to pick the most important spectrum regions. Loading plots for PC1, PC2, and PC3 are shown in Fig. 17.4.

17.3.3 Scree Plots

A Scree plot is a diagnostic tool to check whether the PCA works well on the data. If the most variance is covered by the first two or three PCs, then we can ignore the rest without missing any important information. The eigenvalues at the Y-axis represent the amount of variation. A Scree plot is used to select the principal components to keep. An ideal curve should be steep, then bends at “elbow” point, which is the cut-off point, and flattens out after that. Only PCs 1, 2, and 3 are sufficient to explain the data as suggested by Fig. 17.5.

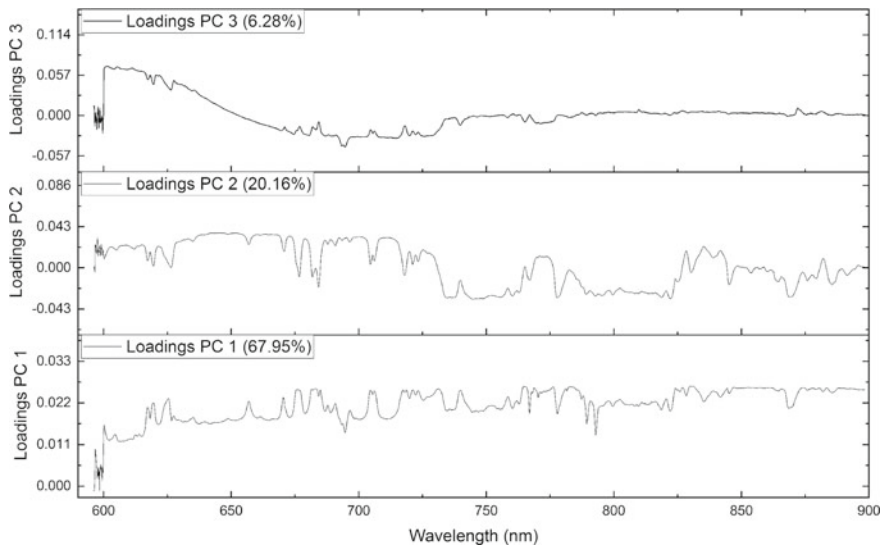
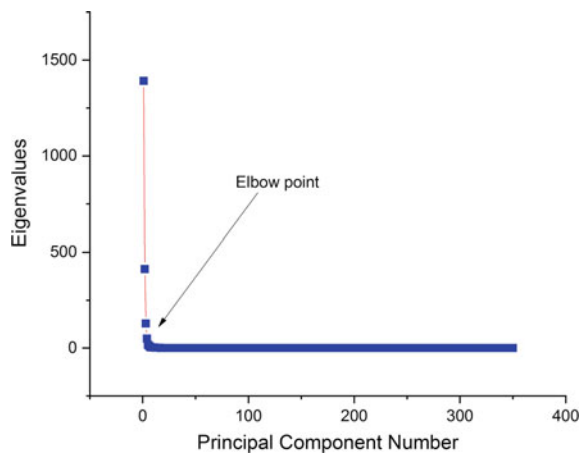


Fig. 17.4 Loadings plot for PC1, PC2 and PC3

Fig. 17.5 Scree plot



17.3.4 Scores Plot

To extract new uncorrelated features called PCs (Principal Components), PCA applies a linear translation to existing spectral features. PCs are sorted in descending order of the variance. PC1 covers the largest variance, PC2 covers the largest variance after PC1, and so on. In this study, we performed PCA on 7 different samples of LIBS spectra. Here, PC1, PC2, and PC3 explained 67.95%, 20.16%, and 6.28% of the data variance respectively counting a total of 94.39% (Fig. 17.6).

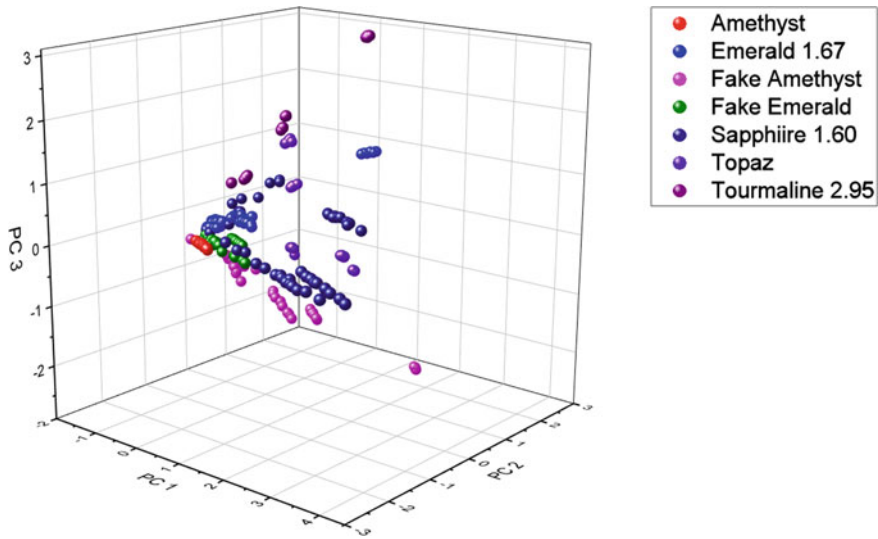


Fig. 17.6 PCA scatter plot

17.3.5 Random Forest

In this experiment, two datasets were used for RF classification. Dataset 1, represents full-length spectra while dataset 2 represents the selected emission lines only, as listed in Table 17.1. In this analysis, 4 splits of training and testing sets were used which are 80% training-20% testing, 70% training-30% testing, 60% training-40% testing, and 50% training-50% testing.

The experiment is repeated 10 times to get the average accuracy to prevent disturbance through randomness. The method from splitting dataset until running the classification accuracy for dataset 2 is the same as the method for dataset 1 but, dataset 2 only chose 14 peaks of an emission line of each gemstones sample. So, with a total of 98 spectra, the classification accuracies for datasets 1 and 2 are summarised in Tables 17.2 and 17.3 respectively.

According to the results, as shown in Tables 17.2 and 17.3, the best accuracy for dataset 1 and dataset 2 is 100% accuracy for 80% training-20% testing split. The

Table 17.1 Selected emission lines for analysis of LIBS spectrum of gemstones (Dataset 2)

Emission line (nm)	Emission line (nm)
619.20	795.44
656.32	799.62
716.18	819.04
742.85	822.14
744.68	824.56
747.28	845.01
777.78	868.50

Table 17.2 Classification accuracies of RF algorithm on Dataset 1

Training test	Testing test	Accuracy (%)									
		1	2	3	4	5	6	7	8	9	10
80	20	100	100	100	100	100	100	100	100	100	100
70	30	95.23	95.23	95.23	95.23	95.23	95.23	95.23	95.23	95.23	95.23
60	40	89.29	89.29	89.29	89.29	89.29	89.29	89.29	89.29	89.29	89.29
50	50	94.29	94.29	94.29	94.29	94.29	94.29	94.29	94.29	94.29	94.29

Table 17.3 Classification accuracies of RF algorithm on Dataset 2

Training test	Testing test	Accuracy (%)									
		1	2	3	4	5	6	7	8	9	10
80	20	100	100	100	100	100	100	100	100	100	100
70	30	90.48	90.48	90.48	90.48	90.48	90.48	90.48	90.48	90.48	90.48
60	40	92.86	92.86	92.86	92.86	92.86	92.86	92.86	92.86	92.86	92.86
50	50	88.57	88.57	88.57	88.57	88.57	88.57	88.57	88.57	88.57	88.57

least accurate dataset 1 was for the 60% training-40% testing set split while dataset 2 shows the least accuracy on 50% training-50% testing split.

17.4 Conclusion

In this study, we have demonstrated LIBS in combination with PCA and RF. PCA was performed on raw spectral data of gemstones to visualize the discriminability of the data and RF helped to classify the gemstones that builds decision trees to accuracy and stable predictions. The results of both training and testing samples demonstrate that the proposed RF model is an effective and efficient approach for the classification of gemstones samples. Prediction accuracy with a perfect 100% score is a strong evidence that LIBS-RF makes a highly potential and exciting tool for the identification gemstones samples.

References

1. S. Domínguez-bella, Gemstones. (2018). <https://doi.org/10.1002/9781119188230.saseas0251>
2. J. Chlachula, Gemstones of eastern Kazakhstan. *Geologos* **26**(2), 139–162 (2020). <https://doi.org/10.2478/logos-2020-0013>
3. S.N. Panya Panya, A.H. Galmed, M. Maaza, B.M. Mothudi, M.A. Harith, J. Kennedy, Laser-induced breakdown spectroscopy (LIBS) on geological samples: Compositional differentiation. *MRS Adv.* **3**(34–35), 1969–1983 (2018). <https://doi.org/10.1557/adv.2018.401>

4. S. Musazzi, U. Perini, *Laser Induced Breakdown Spectroscopy: Theory and Applications*. In Springer Series in Optical Sciences, vol. 182 (2014th ed.). (Springer, 2014)
5. H.X. Yang, H.B. Fu, H.D. Wang, J.W. Jia, M.W. Sigrist, F.Z. Dong, Laser-induced breakdown spectroscopy applied to the characterization of rock by support vector machine combined with principal component analysis. *Chin. Phys. B* **25**(6) (2016). <https://doi.org/10.1088/1674-1056/25/6/065201>
6. T. Zhang, D. Xia, H. Tang, X. Yang, H. Li, Classification of steel samples by laser-induced breakdown spectroscopy and random forest. *Chemom. Intell. Lab. Syst.* **157**, 196–201 (2016). <https://doi.org/10.1016/j.chemolab.2016.07.001>
7. Y. Tian, *Characterization of Laser-induced Plasma and Application to Surface-Assisted LIBS for Powder and Liquid Samples*, 153 (2017)
8. N. Khajehzadeh, T.K. Kauppinen, Fast mineral identification using elemental LIBS technique. *IFAC-PapersOnLine* **28**(17), 119–124 (2015). <https://doi.org/10.1016/j.ifacol.2015.10.089>
9. P. Pořízka, J. Klus, E. Képeš, D. Prochazka, D.W. Hahn, J. Kaiser, On the utilization of principal component analysis in laser-induced breakdown spectroscopy data analysis, a review. *Spectrochim. Acta Part B Atomic Spectrosc.* **148**, 65–82 (May, 2018). <https://doi.org/10.1016/j.sab.2018.05.030>
10. A. Navlani, Attention Required! Cloudflare. Understanding Random Forests Classifiers in Python (16 May 2018)

Chapter 18

Structural Properties of Lithium-Niobium-Calcium Tellurite Glass Doped with Nickel Oxide: Computational Analysis



S. K. Farisya Nadira, H. Nurhafizah, N. H. Ahmad, and Nurul Nadia Adnan

Abstract The structural glass system of lithium-niobium-calcium tellurite glass $\text{TeO}_2 + \text{Li}_2\text{O} + \text{Nb}_2\text{O}_5 + \text{CaO} + \text{NiO}$ was drawn by using GaussView where the bond length and angle need to be set correctly by referring the values on earlier researches. The wavenumber of FTIR spectrum for TeO_2 , Nb_2O_5 , Li_2O , CaO and NiO were 246 cm^{-1} , 885 cm^{-1} , 1127 cm^{-1} , 372 cm^{-1} and 804 cm^{-1} respectively. Highest Raman intensity is obtained in TeO_2 , Nb_2O_5 , Li_2O , CaO and NiO at wavenumber of 942 cm^{-1} , 1012 cm^{-1} , 867 cm^{-1} , 372 cm^{-1} and 804 cm^{-1} respectively. Gaussian error analysis is discussed in this paper as combined structure cannot be optimized and run completely in Gaussian09W. Error such as convergence failure, missing bond and files error are discovered when running whole glass structure. These errors were not found if single molecule is run in Gaussian.

18.1 Introduction

Glass is divided into two categories which are organic and inorganic glass. It is called inorganic because this type of glass consists of inorganic element such as sodium silicates, calcium, potassium, and lead. There was lot of research regarding tellurite glass that has been done. Most of the studies related to tellurite glass were highlighting the doping of this type of glass with rare earth element such as erbium, neodymium and yttrium [1] and not focusing on adding transition metal elements.

S. K. Farisya Nadira · H. Nurhafizah (✉) · N. H. Ahmad
Advanced Optical Material Research Group (AOMRG), Department of Physics, Faculty of
Science, Universiti Teknologi Malaysia, 81310 Bahru, Johor, Malaysia
e-mail: nurhafizah.h@utm.my

N. H. Ahmad
e-mail: nurhidayahahmad@utm.my

N. N. Adnan
Department of Physics and Chemistry, Faculty of Applied Sciences and Technology, Universiti
Tun Hussein Onn Malaysia, 84600 Pagoh, Johor, Malaysia
e-mail: nadia@uthm.edu.my

There are three essential things to be considered when preparing glass which are host, modifiers, and dopant.

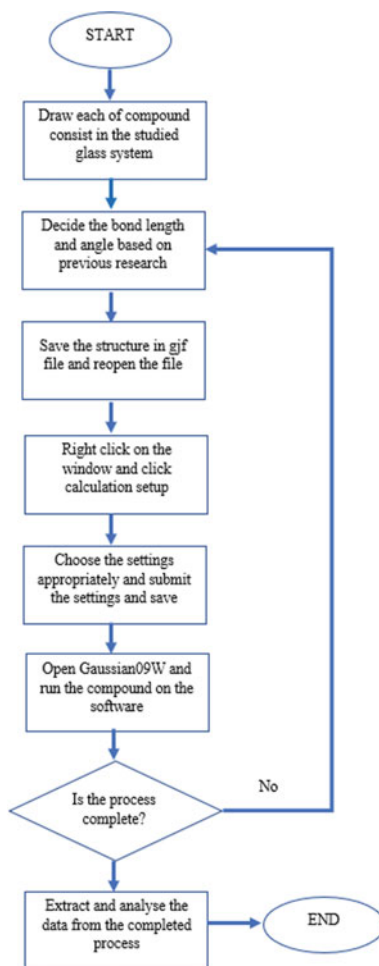
First is host where in this paper tellurite glass is chosen because its unique optical properties and structures [2]. Tellurite glass consists of tellurium dioxide (TeO_2) that has greater stability and high melting point which reached $733\text{ }^\circ\text{C}$ [3]. Tellurite glasses give lot of advantages in science field especially in optics and medical. This type of glass is used in generating, controlling, and detecting photons technologies or in simple word photonics field. Optical amplifiers and micro lenses in networking and pharmaceutical are the example of applications of tellurite glass that has been found in 1990s [4]. Second is modifiers which are commonly involved only small amount of chemicals which will be mixed into glass system to further modify and can enhanced the glass properties for both chemical and physical resulting in improvement on the properties of a glass. The elements used in this paper to act as glass modifier are Li_2O , Nb_2O_5 and CaO . Next, doping is a frequently utilized technique which is the addition of dopant or impurities into other material to intentionally modify both its electrical and optical properties or improving either one because of lack or abundance of electrons [5].

Characterizing of FTIR and Raman spectroscopy was done for each of the molecules in the glass system. FTIR is an analysis which applies mathematical operation which is Fourier transform to extract and interpret the collected reading into the actual spectrum. This method can easily show the absorption and emission of infrared of a sample [1, 5]. Therefore, in this paper the same software will be used to determine the surface chemical composition. Besides, Raman analysis and spectroscopy also used in this research as it shows profoundly specific chemical composition and information related to the sample tested by discovering the crucial vibrational modes of particles. Information about molecular unique characteristics in a system can be provided through the Raman spectra [6].

18.2 Methodology

GaussView application is a software where molecular structure can be sketched easily. From this software, the structures, energies, and vibrational frequency can be known. Figure 18.1 shows the flowchart from sketching to the processing of the structure of $\text{TeO}_2 + \text{Li}_2\text{O} + \text{Nb}_2\text{O}_5 + \text{CaO} + \text{NiO}$ glass system. To sketch the structure, GaussianView is used by selecting the atom of studied glass and combine all the element to form glass system. The bond length and angle need to be set correctly to ensure any further analysis can be done perfectly. After structure is drawn, Gaussian09W is used to run and extract the data of FTIR and Raman spectrum of glass structure. In this paper, only each of glass element is drawn separately and no combination structure as Gaussian failed to optimize the whole glass structure.

Fig. 18.1 Flowchart in drawing and optimizing molecular structures



18.2.1 Molecular Structure

The molecular structure was sketched with actual value of bond length and angle as shown in Fig. 18.2. Then, the drawing was displayed in GaussView and can be analyzed further using FTIR and Raman spectroscopy after optimizing the structure in Gaussian09W. FTIR analysis can be useful when it comes to verify unknown compositions, dopants within polymers, and contamination on a material surface. The FTIR gives the tests' outcome which can precisely tell the details about its molecular configuration. A Raman spectrum presents a few numbers of peaks, displaying the intensity and wavelength range of the Raman scattered light. Every single peak represents to a unique molecular bond vibration which is called as functional groups. Each analysis has different settings on Gaussian calculation which need to be followed to

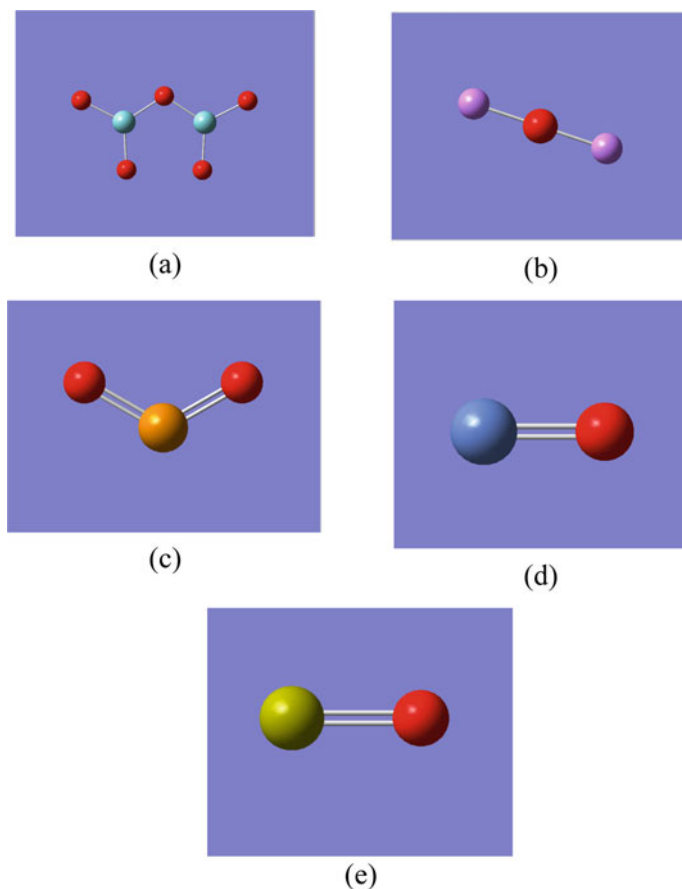


Fig. 18.2 Molecular structure of each molecule consists in glass **a** niobium pentoxide, Nb_2O_5 **b** lithium oxide, Li_2O **c** tellurium dioxide, TeO_2 **d** nickel oxide, NiO **e** calcium oxide, CaO

do computational analysis on the structure since experimental procedure cannot be performed.

18.2.2 Computational Analysis

FTIR analyzed any possible infrared wavelengths region that are absorbed by a material. This is completed by the appliance of infrared (IR) to the samples system. The absorbance of the infrared light's energy by the sample at multiple wavelengths are observed and evaluated to figure out the molecular composition and structure of glass system [7]. Raman spectra define recognizable and separate chemical fingerprint for a particular molecule and can help to verify the material composition within shorter

time or differentiate it from other molecule existed in the material. Peak intensities can be applied to produce quantitative details related to the mixture's configurations [8]. However, since the pandemic cases increases, instead of performing experimental procedure, these analyses were done by applying Gaussian software which the spectrum of each analysis can be acquired from there. The data obtained from software can be plotted directly and analyze by referring the peaks wavelength, wavenumber, and frequency range. Comparison of data obtained with previous experimental values is made to ensure the results are correct.

18.3 Results and Discussion

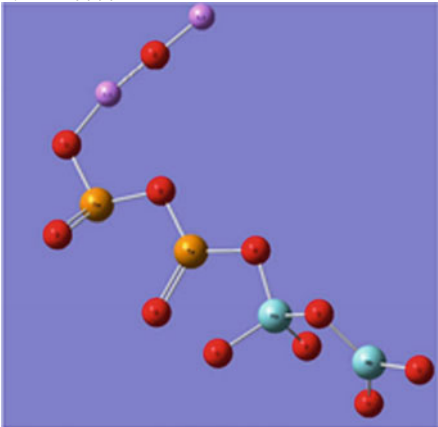
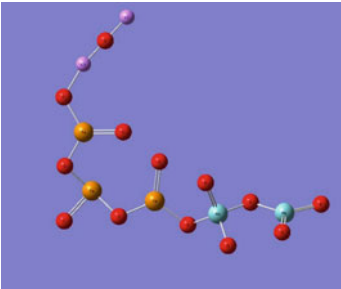
In this section, the results from computational analysis that has been carried out will be presented and compiled. Fourier Transform Infrared (FTIR) and Raman analysis displayed the structure properties of the glass. However, the structure is failed to converge, thus in this section the Gaussian error will be explained in detail which causing the result and all the characterization analysis cannot be obtained.

18.3.1 *Gaussian Error Analysis*

There are three types of errors has been uncovered during this completion of thesis which are Link 9999, Convergence failure-run terminated and Checkpoint files error. These errors were found when whole structure of glass was processed in Gaussian09W. It can be assumed that the structure had wrong connections between bond and incorrect bond length or angle because these are the main factor which led to failure in optimizing the structure [9]. Errors occur when optimization of whole glass system molecular structure is processed in Gaussian09W. Table 18.1 shows type of error might occur when optimizing the molecular structure of glass system.

These errors only occur when whole structure is processed while for single elements, no error was found. The structure can run completely and optimized smoothly. All these errors might be possible to resolve if accurate bond length and perfect connection of all elements can be obtained.

Table 18.1 Types of Gaussian error analysis

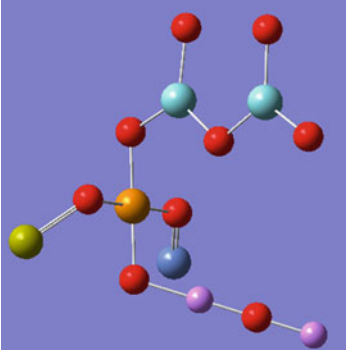
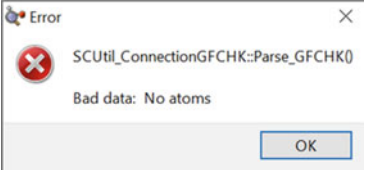
Type of error	Justification
<p>1. Link 9999</p>  <p>Combination of 2TeO₂ + Li₂O + Nb₂O₅ structures</p> <pre> IN TIME YOU MAY BECOME BOSS AND WORK TWELVE HOURS AND HAVE ALL THE WORRY. Error termination request processed by link 9999. Error termination via Lnk1e in C:\Users\FARISYA N </pre>	<p>Structure cannot be optimized, and convergence failed to be achieved. Any further analysis will be incorrect. The most widely recognized factor the error happened is that the geometry optimization is not converged which it can be the length, distance or size of molecules were too far from the real values [9]. That is one of the main reasons that the structure cannot undergo any other analysis as it may lead to wrong results</p>
<p>2. Convergence failure-run terminated</p>  <p>Combination of 3TeO₂ + Li₂O + Nb₂O₅ structures</p> <pre> SCF Done: E(RB3LYP) = -19399.9031383 NFock=128 Conv=0.11D-07 -V. Convergence failure -- run terminated. Error termination via Lnk1e in C:\Users\FAR Job cpu time: 0 days 0 hours 54 min File lengths (MBytes): RWF= 61 Int= </pre>	<p>Output files showed an error after processing the structure. The SCF (self-consistent field) procedure has failed to converge. The SCF procedure might fail to converge if a poor guess is provided for the molecular orbitals. The estimation of bond geometry is inaccurate whether it is too long or too short [10]</p>

(continued)

18.3.2 FTIR and Raman Analysis Spectrum

It is found that for FTIR spectrum range for single element is in between 1200 to 100 cm⁻¹ and for Raman, the wavenumber is obtained in between of 1100 to 373 cm⁻¹ (Table 18.2).

Table 18.1 (continued)

Type of error	Justification
<p>3. Checkpoint files error</p>  <p>Combination of $\text{TeO}_2 + \text{Li}_2\text{O} + \text{Nb}_2\text{O}_5$ structures</p> 	<p>May occur due few conditions which is wrong job type, basis set and method since different structure needs different setting. Wrong setup will directly lead to wrong analysis and wrong interpretation of the studied structure frequency and spectrum. These conditions cause an error which is fchk file errors (bad data). Bad data means the connectivity of the structure in input file is incorrect [11]. The software cannot detect or optimize the structure to allow any further analysis to be done due to the missing bond or bad guess value of the bond. To solve this type of error might request the user to check the bond between each atom whether it is too long or too short. However, in this study, the settings and bond length have been checked and referred to a lot of previous studies yet it still cannot be run and processed</p>

18.3.3 Vibrational Modes of Each Molecule

Section 18.3.3 shows the results on vibrational modes experienced by each molecule that consists in the studied glass. The modes shown in Table 18.3 represented by obvious peaks of wavenumber in the spectrum of intensity versus wavenumber graph displayed in Gaussian software. There are two types of vibrational modes which are stretching and bending vibrational modes. The stretching modes change the angle of molecular bonds which can be divided into few types: scissoring, rocking, wagging, and twisting. While for bending vibrational modes, will undergo the changes in length of the bond between atoms. It can be categorized in two which are: Symmetrical stretching and Asymmetrical stretching.

18.4 Conclusions

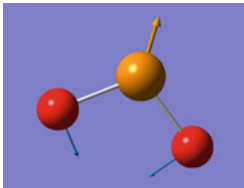
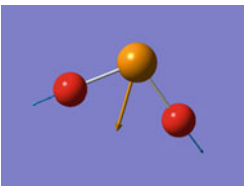
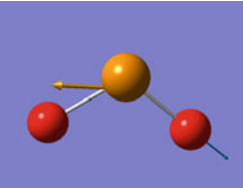
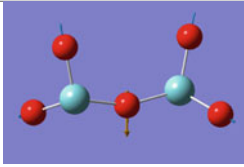
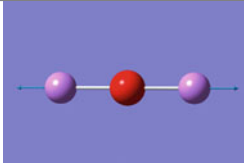
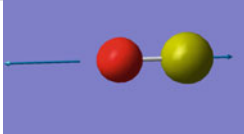
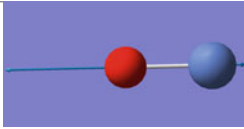
Tellurite glass with composition of $\text{TeO}_2 + \text{Li}_2\text{O} + \text{Nb}_2\text{O}_5 + \text{CaO} + \text{NiO}$ system was analysed using Gaussian09W software and structure was drawn using GaussView. The results obtained throughout this research was led to an uncovering of Gaussian errors that possible to occur when running any structure to be optimized or analysed using FTIR and Raman analysis. It can be concluded that the cause of errors is related

Table 18.2 FTIR and Raman spectrum of all element in studied glass

TIR		Element	Raman	
Wavenumber (cm ⁻¹)	Justification		Wavenumber (cm ⁻¹)	Justification
246 [12]	Not absorbing much of IR light since its wavenumber lower than a range of FTIR [12] Consists of homonuclear diatomic, O ₂ [13, 14]	TeO ₂	247 942 985	Vibrations of Te-O The molecule undergoes bending vibration [19]
885 [15]	Niobium visible in IR region [15]	Nb ₂ O ₅	1012	Vibrations of molecule of Nb-O is in bending modes while for Nb-O-Nb is in stretching vibrational modes
1127	–	Li ₂ O	867	Vibrational mode of bending
372 [16]	Metal oxide can absorb infrared light, low absorbance [16–18]	CaO	372	A stretching vibrational mode Ca-O which changing its bond length by increasing and decreasing of it
804	Same as CaO, low absorbance [16]	NiO	804	Vibrational mode is same as NiO which is stretching vibrations

to the inaccurate bond length between the atoms. Besides, the combined structure of glass cannot be obtained due to Gaussian failure to optimize structure studied even though the setting were corrected and the reference value from previous studies for bond length was used. It can be assumed that the connection of system was done incorrectly. Therefore, this paper decided to discuss about the analysis done only on each/single molecule in TeO₂ + Li₂O + Nb₂O₅ + CaO + NiO glass system. The data of FTIR and Raman analysis were extracted from the Gaussian and plotted spectrum were obtained which in range of 1200–100 cm⁻¹, 1100–373 cm⁻¹ and 200–700 nm respectively.

Table 18.3 Vibrational modes by molecules (*The *blue arrow* represents the displacement vectors while *orange arrow* indicates the dipole derivative unit vector)

Molecules	1st vibrational modes	2nd vibrational modes	3rd vibrational modes
Tellurium dioxide, TeO_2	 <p>Wavenumber: 247 cm^{-1} Scissoring (bending modes)</p>	 <p>Wavenumber: 942 cm^{-1} Symmetrical stretching</p>	 <p>Wavenumber: 985 cm^{-1} Asymmetrical stretching</p>
Niobium pentoxide, Nb_2O_5	 <p>Wavenumber: 1012 cm^{-1} Symmetrical stretching</p>	–	–
Lithium oxide, Li_2O	 <p>Wavenumber: 867 cm^{-1} Symmetric stretching</p>	–	–
Calcium oxide, CaO	 <p>Wavenumber: 372 cm^{-1} Stretching vibration: moving closer and further</p>	–	–
Nickel oxide, NiO	 <p>Wavenumber: 804 cm^{-1} Stretching vibration: moving closer and further</p>	–	–

Acknowledgements We gratefully acknowledge the financial support from Ministry of Higher Education through grant UTMER Vote. Q.J130000.2654.17J86 (Reference No.: PY/2019/01641) of Universiti Teknologi Malaysia, special thanks to the Ministry of Education Malaysia through Fundamental Research Grant Scheme for Research Acculturation of Early Career Researchers (FRGS RACER) (RACER/1/2019/STG02/UTHM//1), Universiti Tun Hussien Onn Malaysia Post-graduate Research Grant (H632) and the experimental assistance from staff of Faculty of Science. The authors are fully acknowledged this institution in completing this paper and helps in making this important research viable and effective.

References

1. N. Hasim, M.S. Rohani, $\text{Er}^{3+}/\text{Nd}^{3+}$ codoped lithium niobate tellurite glass: Improvement in energy transfer induced photoluminescence. *Solid State Phenom. (SSP)*. **290**, 16–21 (2019). <https://doi.org/10.4028/www.scientific.net/SSP.290.16>
2. K.F. Ahmed, Preparation and Characterization of $\text{Nd}^{3+}/\text{Er}^{3+}$ Ions Co-doped Zinc-Tellurite Glass System (July 2016)
3. R.A.H. El-Mallawany, *Tellurite Glasses Handbook: Physical Properties and Data* (2002).
4. L.C. Barbosa, C.O. Filho, E.F. Chillice, Photonic applications of tellurite glasses. *Springer Ser. Mater. Sci.* **254**, 93–100 (2017). https://doi.org/10.1007/978-3-319-53038-3_5
5. Q. Zhou, Z. Fang, J. Li, M. Wang, Applications of TiO_2 nanotube arrays in environmental and energy fields: A review. *Microporous Mesoporous Mater.* **202**(C), 22–35 (2015). <https://doi.org/10.1016/j.micromeso.2014.09.040>
6. C.N. Rinke-Kneapler, M.E. Sigman, *Applications of Laser Spectroscopy in Forensic Science*. Laser Spectroscopy for Sensing: Fundamentals, Techniques and Applications. (Woodhead Publishing Limited, 2014). <https://doi.org/10.1533/9780857098733.3.461>
7. V.A.E. Barrios, J.R.R. Méndez, N.V.P. Aguilar, G.A. Espinosa, J.L.D. Rodríguez, FTIR—An essential characterization technique for polymeric materials. *Infrared Spectrosc. Mater. Sci. Eng. Technol.* (2012). <https://doi.org/10.5772/36044>
8. G. Socrates, Infrared and Raman characteristic group frequencies. Tables and charts. *J. Raman Spectrosc.* (2001). <http://doi.wiley.com/https://doi.org/10.1002/jrs.1238>
9. P.R. Taylor, R. Ahlrichs, *The Structure of F_2O_2 : Theoretical Predictions and Comparisons with F_2 and F_2O* (1982), 72.
10. Z. Sun, T. Zhu, X. Wang, Y. Mei, J.Z.H. Zhang, Optimization of convergence criteria for fragmentation methods. *Chem. Phys. Lett.* **687**, 163–170 (2017). <https://doi.org/10.1016/j.cpl.2017.08.059>
11. J. Barroso, fchk file errors (Gaussian) missing or bad Data: RBondIDr. Joaquin Barroso's Blog. (14 Dec 2017). <https://joaquinbarroso.com/2017/12/14/fchk-file-errors-gaussian-missing-or-bad-data-rbond/>
12. P. Mohamed Shameer, P. Mohamed Nishath, *Exploration and Enhancement on Fuel Stability of Biodiesel: A Step Forward in the Track of Global Commercialization*. In *Advanced Biofuels: Applications, Technologies and Environmental Sustainability*. (Elsevier Ltd, 2019). <https://doi.org/10.1016/B978-0-08-102791-2.00008-8>
13. J. Lacosse, *Real Time Emission Measurements Using FTIR Spectroscopy (EPA Method 320)* (2010). www.spectralinsights.com
14. V. Rodriguez, M. Couzi, F. Adamietz, M. Dussauze, G. Guery, T. Cardinal, P. Veber, K. Richardson, P. Thomas, Hyper-Raman and Raman scattering in paratellurite TeO_2 . *J. Raman Spectrosc.* **44**(5), 739–745 (2013). <https://doi.org/10.1002/jrs.4251>
15. T. Ikaya, M. Senna, Change in the structure of niobium pentoxide due to mechanical and thermal treatments. *J. Non-Cryst. Solids* **105**(3), 243–250 (1988). [https://doi.org/10.1016/0022-3093\(88\)90313-4](https://doi.org/10.1016/0022-3093(88)90313-4)

16. Ō Kammori, N. Yamaguchi, K. Sato, Infrared absorption spectra of metal oxides. *Bunseki Kagaku* **16**(10), 1050–1055 (1967). <https://doi.org/10.2116/bunsekikagaku.16.1050>
17. N.T. McDevitt, W.L. Baun, Infrared absorption study of metal oxides in the low frequency region (700–240 cm^{-1}). *Spectrochim. Acta* **20**(5), 799–808 (1964). [https://doi.org/10.1016/0371-1951\(64\)80079-5](https://doi.org/10.1016/0371-1951(64)80079-5)
18. R. Mohadi, K. Anggraini, F. Riyanti, A. Lesbani, Preparation calcium oxide from chicken eggshells. *Sriwijaya J. Environ.* **1**(2), 32–35. (2016). <https://doi.org/10.22135/sje.2016.1.2.32-35>
19. T. Myint, S. Soe Thin, P. Kaung, S. Htoon, Infrared spectroscopy and Raman scattering studies on the structure of Ag_2O . B_2O_3 . TeO_2 glass. *J. Myan. Acad. Arts Sci* **IV**(2) (2006)

Part II
Mathematics

Chapter 19

Comparing the Hamming and Manhattan Heuristics in Solving the 8—Puzzle by A* Algorithm



Siauw Chong Lee  and Tyan Her See

Abstract The 8-puzzle is a classical sliding puzzle that consists of 8 numbered tiles placed in random order on a board with 9 cells. The goal is to rearrange the tiles by sliding them horizontally or vertically into the vacant cell such that the number in each tile is ordered in ascending order. 8-puzzles can be solved by using the path searching algorithms such as Best-first search, Depth-first search, Dijkstra's algorithm, or the A* algorithm. In this paper, the A* algorithm will be the one chosen. The A* algorithm is known to be an improvement to the Dijkstra's algorithm since not all potential paths need to be searched before the optimal one is found. The key of the A* algorithm is choosing the right heuristic formula as an estimation to the cost of each path. Two heuristics, the Hamming distance and the Manhattan distance are chosen to be implemented in the A* algorithm. The computing time for our several test puzzles suggests that the Manhattan distance is found to be the more efficient heuristic.

19.1 Introduction

N -puzzles consist of N numbers where N can be 8, 15, 24, and so on. Such values are picked because the puzzle is played on a square board of m^2 cells ($m > 2$) where $N = m^2 - 1$ [1]. An N -puzzle board must have one blank space to allow the numbers to move within the board and the goal is simple; slide the numbers to the vacant spot until the numbers are ascendingly ordered. An example of N -puzzle, the 8-puzzle, will be solved in this paper.

Not all N -puzzles are solvable. If we start from a solved puzzle and simply swap two numbers, there are situations where we could not return the numbers to their original position by sliding them. In 1891, Sam Loyd [2] posted a problem which was known as Sam's 14–15 puzzle and the puzzle was impossible to solve because

S. C. Lee (✉) · T. H. See

Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, Pagoh Edu Hub, 84800 Parit Raja, Johor, Malaysia

e-mail: sclee@uthm.edu.my

all the numbers are in their correct position except for two; the numbers 14 and 15 were reversed. Sam Loyd offers a \$1000 reward for the first correct solution for this problem, but the reward was never claimed.

Many pathfinding techniques can find the least moves needed to solve the puzzle. Methods like Dijkstra's algorithm can be implemented but the A* algorithm will be the pathfinding technique of choice in this paper. The Dijkstra's algorithm, introduced by Edsger W. Dijkstra in 1956 is a pathfinding algorithm that finds the optimal path between nodes in a graph. Starting from a node, the algorithm chooses the node that yields the least cost out of the possible adjacent nodes until all nodes are considered. The A* algorithm, on the other hand, is an extension of the Dijkstra's algorithm [3] by considering the additional information, known as the heuristics in searching the next best node. The heuristics are functions that estimates how far is the cost of the solution from the current state, the better the heuristics, the faster the A* algorithm finds the optimal path. The heuristics must be carefully chosen though, a heuristic that underestimates the actual cost too much might result in longer computing time, while an overestimating heuristic might not find the optimal path at all. An admissible heuristic is the heuristic that never overestimates the actual cost and the admissible heuristics that we will use in solving the 8-puzzle are the Manhattan distance and the Hamming distance. We compared the computational time for each of the heuristics using 3 sets of solvable puzzles and found out that the Manhattan heuristic outperforms the Hamming heuristic.

19.2 A* Algorithm

A* algorithm can guarantee to find the shortest path like Dijkstra's algorithm if an admissible heuristic is used. This algorithm will choose the next node by considering the cost of getting to that node. The A* algorithm is one of the best methods for solving the sliding puzzle [4]. Before solving the 8-puzzle with the A* algorithm, we should make sure that the puzzle is solvable, and this can be achieved by the rules stated in Sect. 19.2.1.

19.2.1 Solvability of an N-puzzle

There are $9!$ (362, 880) possible initial states for an 8-puzzle but only $\frac{9!}{2}$ (181, 440) initial positions can lead to the goal state [5]. That means only half of the random initial positions can be solved. The solvability of a puzzle can be determined without solving the puzzle itself by calculating the number of inversions for that puzzle. An inversion happened when a tile precedes other tiles with a lower number on it. In other words, inversion between a pair of tiles (a, b) is a occurs before b , but $a > b$. The three simple rules to test whether an N -puzzle is solvable or not are as follows.

1. If the puzzle size is odd, then the puzzle can be solved if the inversion number is even in the initial state.
2. If the size is even, then the puzzle can be solved if there are an even number of inversions and if the blank cell is on an odd-numbered row counting from the bottom.
3. If the size is even, then the puzzle can be solved if there are an odd number of inversions and if the blank cell is on an even-numbered row counting from the bottom.

19.2.2 Solving the 8-puzzle by A* Algorithm

Consider the 8-puzzle in Fig. 19.1. Starting from the initial state at the top, there are 2 possible moves; we can move 3 up or 2 to the right. Each of these will result in a new state and each state can be considered as a node and a path is formed by connecting a series of nodes to the one adjacent to it. The number of moves taken to reach a certain state is considered as the cost of the path taken to reach the node

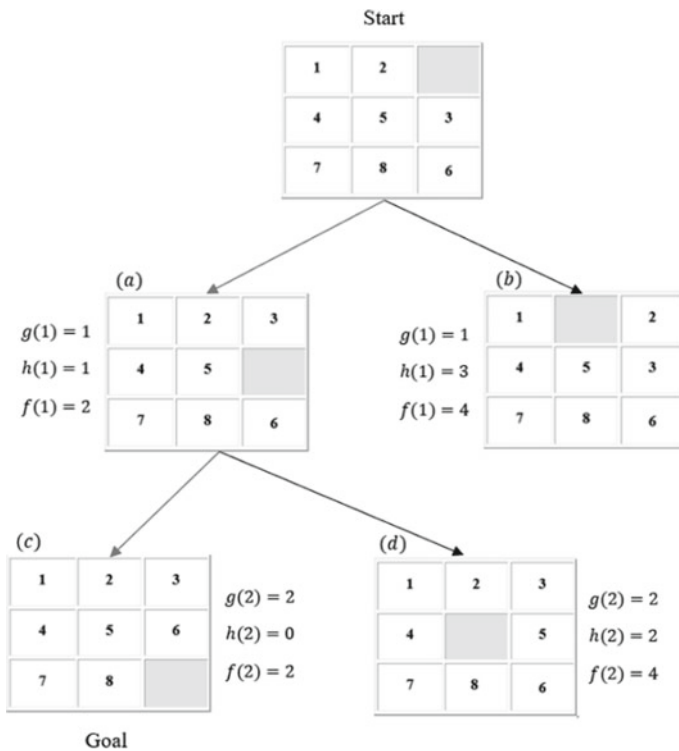


Fig. 19.1 Example of solving 8-puzzle by A* algorithm

that corresponds to that state. The optimal path, which is the one we intend to find, is the solution that takes the least moves. In other words, it is the path that cost the least from the initial node to the goal node.

The A* algorithm is essentially the best first search method implemented with the following function:

$$f(n) + g(n) = h(n) \quad (19.1)$$

where,

$g(n)$ = the cost from the starting node to the current node.

$h(n)$ = heuristic value, the estimated cost from the current node to the goal node

n = n th node on the path.

The two possible types of functions that h in (1) are.

1. The Hamming distance: the total tiles out of its goal position.
2. The Manhattan distance: the sum of the shortest path of each tile to its goal position [6].

The A* algorithm manages the path searching by constantly updating two lists, the open list and the closed list. The open list contains the unexplored nodes while the closed list contains the explored ones. Referring to Fig. 19.1, the starting node is set as the current node. Now there are 2 possible moves, which results in 2 adjacent nodes (a) and (b). Both nodes were never visited before, so we place them in the open list. The next thing to do is calculate the f , g and h values for both nodes, which are shown in Fig. 19.1.

Since the f value for node (a) is smaller than the f value for node (b), we choose node (a) as the current node. Now starting node is already examined, it is placed in the closed list. From the current node (node (a)), we find the nodes adjacent to it, which are nodes (c) and (d) and add them to the open list. Again, we calculate the f , g and h values for both nodes (as shown in Fig. 19.1) and node (c) is found to be optimal between those two nodes. Since we had reached the goal node, the search is done, and the optimal path is Start → (a) → (b). If (c) is not the goal node, it will be selected as the current node and (a) will be added to the closed list. The same process is repeated until the goal node is reached [7]. In Fig. 19.1, both Hamming and Manhattan heuristic yield the same value, coincidentally. The flowchart of the whole process is illustrated in Fig. 19.2.

19.3 Experiments and Results

For each heuristic, we record the running time for 10 trials with the same puzzle and take the average as the benchmark time. The A* algorithm is programmed in Java and all experiments are done in a windows system with AMD Athlon 200GE processor and 8 GB RAM. Figure 19.3 shows the test puzzles used in our experiments.

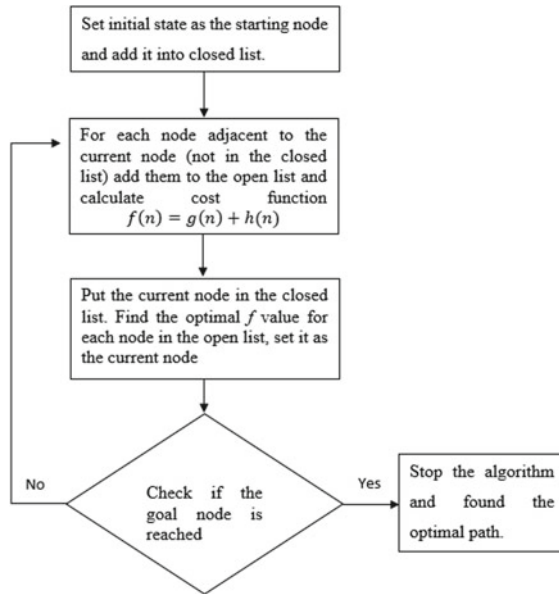


Fig. 19.2 Flow chart of the A* algorithm

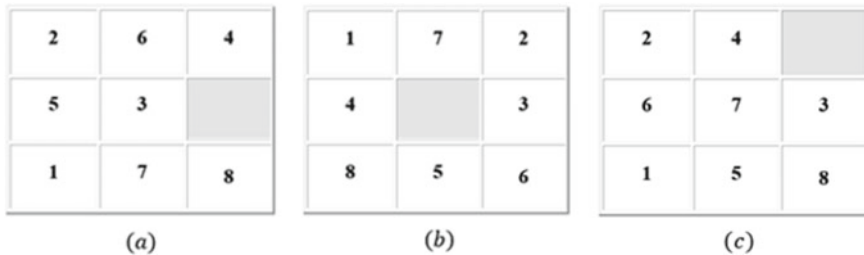


Fig. 19.3 Test puzzles

In Table 19.1, the time taken to solve the puzzle (a), (b), and (c) with the Manhattan heuristic were 22.5, 42 and 37.5 ms respectively whereas Hamming heuristic clocked 178.7, 136.3 and 900.8 ms, respectively. From these experiments we can see that the Manhattan heuristic clearly outperforms the Hamming heuristic. A possible reason is the Manhattan distance provides a much more accurate estimation to the moves

Table 19.1 Computational time of A* using Manhattan distance and Hamming distance

Puzzle	Manhattan distance (ms)	Hamming distance (ms)
(a)	22.5	178.7
(b)	42	136.3
(c)	37.5	900.8

needed to reach the goal than the Hamming distance. In our experiment, we observe that the Hamming heuristic often gives the same optimal h value across many nodes in the open list, whereas for the Manhattan heuristic, the h values are different from one node to another. This situation is depicted in Fig. 19.4. From the initial node, the three possible moves are shown in Fig. 19.4a–c. The Hamming distances are the same for these three nodes, but the Manhattan distance suggests that node (a) should be the one chosen. Since all nodes are optimal for Hamming heuristic, choosing one

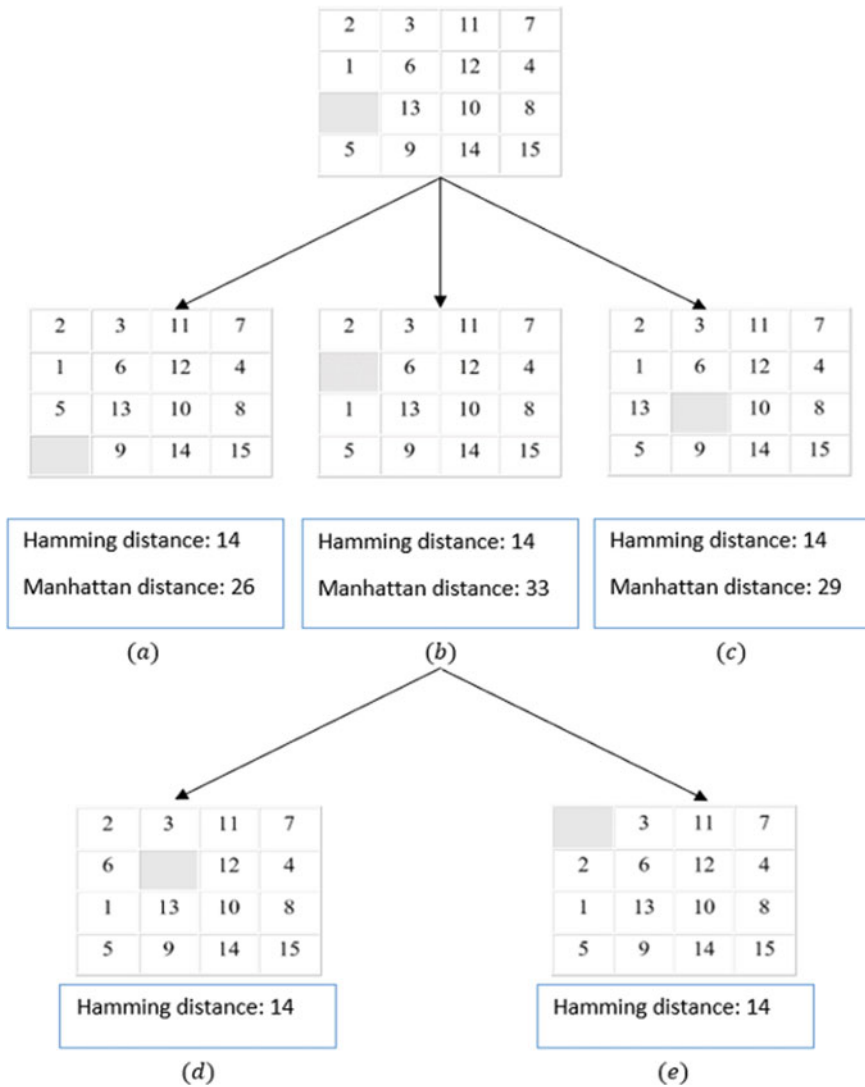


Fig. 19.4 Hamming distance and Manhattan distance of a particular step in a 15-puzzle

of them arbitrarily, in this case let's say if (b) is selected, again yields the similar situation as before, same optimal Hamming distance for all nodes ((d) and (e)). Not only this will create more branches, but also slows down the searching progress and thus cost more computational time than the Manhattan heuristic.

19.4 Conclusion

We conclude that when solving the 8-puzzle with the A* algorithm, the Manhattan heuristic is much efficient than the Hamming heuristic and the reason is the actual cost from a node to its goal is much accurately estimated by the Manhattan distance than the Hamming distance.

Reference

1. E.O. Osaghae, An alternative solution to N-puzzle problem. *J. Appl. Sci. Environ. Manag.* **22**(8), 1199–1205 (2018). <https://doi.org/10.4314/jasem.v22i8.9>
2. A.K. Mishra, P.C. Siddalingaswamy, Analysis of Tree Based Search Techniques for Solving 8-puzzle Problem. In *Conference 2016, LNCS 9999* (Springer Heidelberg, 2016), pp. 1–13
3. R.E. Korf, Recent Progress in the Design and Analysis of Admissible Heuristic Functions. In *International Symposium on Abstraction, Reformulation, and Approximation*. (Springer, Berlin, 2000), pp. 45–55. https://doi.org/10.1007/3-540-44914-0_3
4. B.M. Sathyaraj, L.C. Jain, A. Finn, S. Drake, Multiple UAVs path planning algorithms: A comparative study. *Fuzzy Optim. Decis. Making* **7**(3), 257–267 (2008). <https://doi.org/10.1007/s10700-008-9035-0>
5. K. Mathew, M. Tabassum, Experimental comparison of uninformed and heuristic AI algorithms for N puzzle and 8 queen puzzle solution. *Int. J. Digital Inf. Wirel. Commun.* **4**(1), 143–154 (2014). <https://doi.org/10.17781/P001092>
6. A. Reinefeld, Complete Solution of the Eight-Puzzle and the Benefit of Node Ordering in IDA*. In *Proceedings of the 13th International Joint Conference on Artificial Intelligence*, vol. 1 (1993), pp. 248–253
7. R.E. Korf, A. Felner, Disjoint pattern database heuristics. *Artif. Intell.* **134**(2), 9–22 (2002). [https://doi.org/10.1016/S0004-3702\(01\)00092-3](https://doi.org/10.1016/S0004-3702(01)00092-3)

Chapter 20

Magnetohydrodynamic Effects in Mixed Convection Viscoelastic Fluid Flow at Lower Stagnation Point on a Sliced Sphere



Basuki Widodo, Tunjung Eviloka, Chairul Imron, and Dieky Adzkiya

Abstract In this paper, mixed convection of viscoelastic fluid at the lower stagnation point on a magnetic sliced sphere is investigated. Dimensional governing equation and boundary condition of the problem are further developed. The dimensional governing equation and the boundary condition are further transformed into non-dimensional form using non-dimensional variables. We further obtain non-linear partial differential equation. The non-linear partial differential equation and non-similarity transformation with the consideration of the effect of magnetohydrodynamic (MHD) are solved numerically by using Keller-box method. The influences of sliced sphere magnetic field on viscoelastic fluid flow and heat transfer are then discussed. The results depicted that the uniform magnetic field produced by Lorentz force and slicing on the sphere act as determining factors for the trend of viscoelastic fluid movement and controlling the cooling rate of the sphere surface.

Keywords Sliced sphere · Viscoelastic fluid · Mixed convection

20.1 Introduction

Fluid is a substance that is able to deform continuously when subjected to shear stress, although the shear stress is quite small. Fluid that has a constant viscosity and linear stress curve in the presence of a working force is called Newtonian fluid. However, for fluids that have properties other than that are called Non-Newtonian fluids [1]. In many industrial sectors, fluid heating or cooling is widely used [2].

The study of Magnetohydrodynamics (MHD) has recently been very interesting and has been widely carried out. This magnetohydrodynamic fluid flow is a fluid flow that is influenced by a magnetic field and can conduct electric current [3]. The use of magnetohydrodynamics in the field of technology is quite extensive. One example is

B. Widodo (✉) · T. Eviloka · C. Imron · D. Adzkiya
Department of Mathematics, Institut Teknologi Sepuluh Nopember, Kampus ITS Sukolilo,
Surabaya 60111, Indonesia
e-mail: b_widodo@matematika.its.ac.id

in magnetohydrodynamic power plants. By applying the MHD principle in the power plant, the efficiency can reach 30% [4]. Another example of the implementation of this MHD is also in oil drilling [5].

Viscoelastic fluid is one type of Non-Newtonian fluids besides pseudoplastic fluid, rheopectic fluid, and thixotropic fluid which has viscous and elastic properties. This viscoelastic fluid has attracted the attention of many researchers because the application of this fluid is very important [6]. Such as Widodo et al. [7] who have investigated the effect of viscoelastic fluid flow passing through a flat plate. Sahaya et al. [6] have also conducted research on the flow of suspended viscoelastic magnetohydrodynamic fluid passing through a flat plate. This viscoelastic fluid has been widely studied under the influence of various types of convection and passes through various types of objects such as elliptical cylinders, porous spheres, flat plates and solid spheres.

In this paper, we discuss mixed convection of viscoelastic fluid at the lower stagnation point on a magnetic sliced sphere. We develop dimensional governing equations and their boundary conditions based on the laws of conservation of mass, momentum, and energy. The viscoelastic fluid flows from the bottom up and passes through the sliced magnetic solid sphere. The dimensional governing equations are further transformed into non-dimensional governing equations using non-dimensional variables. Then obtained a non-linear partial differential equation. Furthermore, this equation is converted into non-linear ordinary differential equation using the stream function and similarity equations. The non-linear ordinary differential equations are then solved numerically using the Keller-Box method. The numerical results obtained are in the form of temperature and fluid velocity profiles by varying the values of the magnetic parameters, convection parameters, viscoelastic parameters, Prandtl numbers, and spherical slice angles.

20.2 Research Method

We consider the viscoelastic fluid flow that passes through a vertical porous cylinder with the upstream velocity U_∞ , upstream velocity temperature T_∞ , surface temperature T_w , slice angle θ_s , gravity force g and radius of sphere a . Physical model of the problem and the coordinate system used in solving the problem can be seen in Fig. 20.1a, b, respectively.

20.2.1 Developing Dimensional Governing Equations

The dimensional governing equations are developed from dimensional equations that are obtained from the conservations of mass, momentum and energy. We use the second Newton Law to obtain the momentum equation, and using the first conservative of thermodynamics to obtain the energy equation. We further obtain the dimensional governing equation as follows:

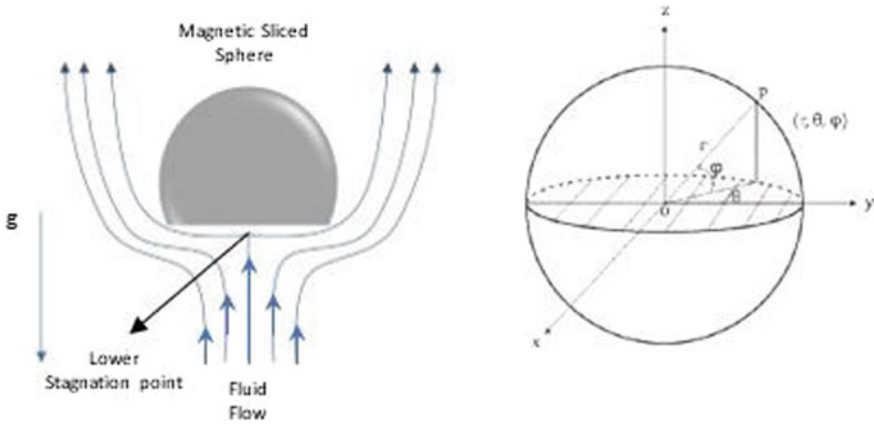


Fig. 20.1 The left panel represents the physical model of the problem, whereas the right panel represents the coordinate system

1. Continuity equation

$$\frac{\partial(\bar{r}\bar{u})}{\partial\bar{x}} + \frac{\partial(\bar{r}\bar{v})}{\partial\bar{y}} = 0 \tag{20.1}$$

2. Momentum equation

a. In x -axis

$$\begin{aligned} \left(\bar{u}\frac{\partial\bar{u}}{\partial\bar{x}} + \bar{v}\frac{\partial\bar{u}}{\partial\bar{y}}\right) &= -\frac{1}{\rho}\frac{\partial p}{\partial\bar{x}} + \frac{\mu_0}{\rho}\left(\frac{\partial^2\bar{u}}{\partial\bar{x}^2} + \frac{\partial^2\bar{u}}{\partial\bar{y}^2}\right) - \frac{k_0}{\rho}\left[\bar{u}\left(\frac{\partial^3\bar{u}}{\partial\bar{x}^3} + \frac{\partial^3\bar{u}}{\partial\bar{x}\partial\bar{y}^2}\right)\right. \\ &+ \bar{v}\left(\frac{\partial^3\bar{u}}{\partial\bar{x}^2\partial\bar{y}} + \frac{\partial^3\bar{u}}{\partial\bar{y}^3}\right) - \frac{\partial\bar{u}}{\partial\bar{y}}\left(\frac{\partial^2\bar{u}}{\partial\bar{x}\partial\bar{y}} + \frac{\partial^2\bar{v}}{\partial\bar{x}^2}\right) - 2\frac{\partial\bar{v}}{\partial\bar{x}}\frac{\partial^2\bar{u}}{\partial\bar{x}\partial\bar{y}} \\ &\left. - \frac{\partial\bar{u}}{\partial\bar{x}}\left(3\frac{\partial^2\bar{u}}{\partial\bar{x}^2} - \frac{\partial^2\bar{u}}{\partial\bar{y}^2}\right)\right] - \frac{1}{\rho}\frac{25}{16}\sigma B_0^2\bar{u} + g\beta(\bar{T} - T_\infty)\tan\left(\frac{x\cos x}{\cos\theta_s}\right) \end{aligned} \tag{20.2}$$

b. In y -axis

$$\begin{aligned} \left(\bar{u}\frac{\partial\bar{v}}{\partial\bar{x}} + \bar{v}\frac{\partial\bar{v}}{\partial\bar{y}}\right) &= -\frac{1}{\rho}\frac{\partial p}{\partial\bar{y}} + \frac{\mu_0}{\rho}\left(\frac{\partial^2\bar{v}}{\partial\bar{x}^2} + \frac{\partial^2\bar{v}}{\partial\bar{y}^2}\right) - \frac{k_0}{\rho}\left[\bar{u}\frac{\partial^3\bar{v}}{\partial\bar{x}^3} + \bar{u}\frac{\partial^3\bar{v}}{\partial\bar{x}\partial\bar{y}^2}\right. \\ &+ \bar{v}\frac{\partial^3\bar{v}}{\partial\bar{x}^2\partial\bar{y}} + \bar{v}\frac{\partial^3\bar{v}}{\partial\bar{y}^3} + \frac{\partial\bar{u}}{\partial\bar{x}}\left(3\frac{\partial^2\bar{v}}{\partial\bar{y}^2} - \frac{\partial^2\bar{v}}{\partial\bar{x}^2}\right) - \frac{\partial\bar{v}}{\partial\bar{x}}\left(\frac{\partial^2\bar{v}}{\partial\bar{x}\partial\bar{y}} + \frac{\partial^2\bar{u}}{\partial\bar{y}^2}\right) \\ &\left. - 2\frac{\partial\bar{u}}{\partial\bar{y}}\frac{\partial^2\bar{v}}{\partial\bar{x}\partial\bar{y}}\right] - \frac{1}{\rho}\frac{25}{16}\sigma B_0^2\bar{v} + g\beta(\bar{T} - T_\infty)\frac{g}{\cos\left(\frac{x\cos x}{\cos\theta_s}\right)} \end{aligned} \tag{20.3}$$

3. Energy equation

$$\bar{u} \frac{\partial \bar{T}}{\partial \bar{x}} + \bar{v} \frac{\partial \bar{T}}{\partial \bar{y}} = \alpha \frac{\partial^2 \bar{T}}{\partial \bar{y}^2} \tag{20.4}$$

with boundary conditions

$$\begin{aligned} \bar{u} = \bar{v} = 0, \quad \bar{T} = T_w & \quad \text{for } \bar{y} = 0 \\ \bar{u} = \bar{u}_e(x), \quad \frac{\partial \bar{u}}{\partial \bar{y}} = 0, \quad \bar{T} = T_\infty & \quad \text{for } \bar{y} \rightarrow \infty \end{aligned}$$

where free flow velocity $\bar{u}_e = \frac{3}{2} \sin\left(\frac{\bar{x}}{b}\right)$ and α is the thermal diffusion, μ_0 is the dynamic viscosity, k_0 is a coefficient, ρ is the density of fluid, β is the coefficient of thermal expansion, \bar{u} is the component of velocity on the \bar{x} axis, and \bar{v} is the component of velocity on the \bar{y} axis.

20.2.2 Boundary Layer Approach

The Equations (20.2) and (20.3) are simplified by using boundary layer theory. The simplification process can be seen in the reference source [8–14]. Furthermore, The Equations (20.2) and (20.3) are reviewed in the form 1 and Δ , in which 1 indicates the dominant fluid velocity and Δ indicates a very small change in fluid velocity [15].

$$\bar{u} \sim 1, \quad \bar{x} \sim 1, \quad \bar{v} \sim \Delta, \quad \frac{K_0}{\rho} \sim \Delta^2, \quad \frac{\mu_0}{\rho} \sim \Delta^2, \quad g \sim 1$$

Based on the form of notation above, the process of simplification of momentum equations on the x and y axis direction can be written as follows. In x -axis:

$$\begin{aligned} \left(\bar{u} \frac{\partial \bar{u}}{\partial \bar{x}} + \bar{v} \frac{\partial \bar{u}}{\partial \bar{y}} \right) = -\frac{1}{\rho} \frac{\partial p}{\partial \bar{x}} + \nu \left(\frac{\partial^2 \bar{u}}{\partial \bar{y}^2} \right) - \frac{k_0}{\rho} \left[u \left(\frac{\partial^3 \bar{u}}{\partial \bar{x} \partial \bar{y}^2} \right) + \nu \left(\frac{\partial^3 \bar{u}}{\partial \bar{y}^3} \right) \right. \\ \left. - \frac{\partial \bar{u}}{\partial \bar{y}} \left(\frac{\partial^2 \bar{u}}{\partial \bar{x} \partial \bar{y}} \right) + \frac{\partial \bar{u}}{\partial \bar{x}} \frac{\partial^2 \bar{u}}{\partial \bar{y}^2} \right] - \frac{25}{16} \frac{1}{\rho} \sigma B_0^2 \bar{u} + g\beta(\bar{T} - T_\infty) \tan\left(\frac{x \cos x}{\cos \theta_s}\right) \tag{20.5} \end{aligned}$$

while in y -axis direction, the result contains only the value of the Δ . This indicates that the momentum equation on the y -axis direction undergoes a very small change in velocity. Thus, in the next analysis the momentum equation on the y -axis is not considered.

20.2.3 Non-dimensional Governing Equations

We introduce non-dimensional variables as follows [11]:

$$x = \frac{\bar{x}}{a}, \quad y = Re^{\frac{1}{2}} \frac{\bar{y}}{a}, \quad r = \frac{\bar{r}}{a}, \quad b = \frac{\bar{b}}{a}, \quad u = \frac{\bar{u}}{U_{\infty}}$$

$$v = Re^{\frac{1}{2}} \frac{\bar{v}}{U_{\infty}}, \quad p = \frac{\bar{p}}{\rho U_{\infty}^2}, \quad T = \frac{\bar{T} - T_{\infty}}{T_w - T_{\infty}}$$

with Reynolds number $Re = \frac{U_{\infty} a}{\nu}$ where ν is the kinematic viscosity written as $\nu = \frac{\mu}{\rho}$.

We also use the following parameters

- Magnetic parameter (M):

$$M = \frac{\sigma B_0^2 a}{\rho U_{\infty}}$$

- Prandtl number (Pr):

$$Pr = \frac{\nu}{\alpha}$$

- Viscoelastic parameter (K):

$$K = \frac{k_0 U_{\infty}}{a \rho \nu}$$

- Convection parameter (λ):

$$\lambda = \frac{Gr}{Re^2}$$

where the Grashof number is $Gr = g\beta(T - T_{\infty})\frac{a^3}{\nu^2}$. Then convection parameter λ can be written as

$$\lambda = \frac{g\beta(T - T_{\infty})a^3}{U_{\infty}^2}$$

By substituting the non-dimensional variables and non-dimensional parameters mentioned above to the non-dimensional governing equations, we obtain:

1. Continuity equation

$$\frac{\partial(ru)}{\partial x} + \frac{\partial(rv)}{\partial y} = 0 \quad (20.6)$$

2. Momentum equation

$$\frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} = -\frac{\partial p}{\partial x} + \frac{\partial^2 u}{\partial y^2} - K \left[\frac{\partial}{\partial x} \left(u \frac{\partial^2 u}{\partial y^2} \right) + v \frac{\partial^3 u}{\partial y^3} - \frac{\partial u}{\partial y} \frac{\partial^2 u}{\partial x \partial y} \right] - \frac{25}{16} Mu + \lambda T \tan \left(\frac{x \cos x}{\cos \theta_s} \right) \tag{20.7}$$

3. Energy equation

$$u \frac{\partial T}{\partial x} + v \frac{\partial T}{\partial y} = \frac{1}{Pr} \frac{\partial^2 T}{\partial y^2} \tag{20.8}$$

with the following boundary conditions:

$$u = v = 0, \quad T = 1 \quad \text{for } y = 0$$

$$u = u_e(x), \quad \frac{\partial u}{\partial y} = 0, \quad T = 0 \quad \text{for } y \rightarrow \infty$$

20.2.4 Develop Non-similarity Equations

To change the non-linear partial differential equations to non-linear ordinary differential equations, we introduce the stream function [12]:

$$\psi = xr(x)f(x, \eta), \quad T = \theta(x, \eta)$$

and ψ is a flow function that is defined as follows:

$$u = \frac{1}{r} \frac{\partial \psi}{\partial \eta}$$

$$v = -\frac{1}{r} \frac{\partial \psi}{\partial x}$$

By substituting the stream function into (20.6)-(20.8), and applying the similarity equations, we obtain the non-linear ordinary differential equations as follows:

1. Continuity equation

$$\frac{\partial^2 \psi}{\partial x \partial \eta} - \frac{\partial^2 \psi}{\partial x \partial \eta} = 0$$

Because the continuity equation is 0, then the continuity equation is trivial so it is not used in subsequent calculations.

2. Momentum equation

$$f''' + 2ff'' - (f')^2 + \frac{9}{4 \cos^2 \theta_s} - 2K(f'f''' - ff'''' - (f'')^2) + \frac{\lambda\theta}{\cos \theta_s} - \frac{25}{16}M\left(f' - \frac{3}{2 \cos \theta_s}\right) = 0 \tag{20.9}$$

3. Energy equation

$$\frac{1}{Pr}\theta'' + 2f\theta' = 0 \tag{20.10}$$

with the boundary conditions

$$f(0) = f'(0) = 0, \quad \theta = 1 \quad 0$$

$$f'(\infty) = 1, \quad f''(\infty) = 0, \theta(\infty) = 0 \quad \eta \rightarrow \infty$$

with sign ' denotes a derivative with respect to η .

20.2.5 Numerical Solution

We use Keller-Box method to solve (20.9) and (20.10). The steps are as follows:

1. Transformation into first order By defining the following function:

$$f' = u$$

$$u' = v$$

$$v' = w$$

$$s' = t$$

So that (20.9)-(20.10) can be written as follows:

$$w + 2fv - u^2 + \frac{9}{4 \cos^2 \theta_s} - 2K(uw - fw - v^2) + \frac{\lambda s}{\cos \theta_s} - \frac{25}{16}M\left(u - \frac{3}{2 \cos \theta_s}\right) = 0 \tag{20.11}$$

$$\frac{1}{Pr}t' + 2ft = 0 \tag{20.12}$$

with the following boundary conditions:

$$f(0) = u(0) = 0, \quad s(0) = 0 \quad \text{when } \eta = 0$$

$$u \rightarrow 1, \quad s \rightarrow 0 \quad \text{when } \eta \rightarrow \infty$$

2. Discretization of the equation

Furthermore, discretization is carried out on the obtained equations by using central difference methods.

3. Linearization of the equations

After obtaining the results of the discretization of the equations, then linearization is then conducted with the form of iteration (*i*) using Newton’s methods, as follows:

$$f_j^{(i+1)} = f_j^{(i)} + \delta f_j^{(i)}$$

$$u_j^{(i+1)} = u_j^{(i)} + \delta u_j^{(i)}$$

$$v_j^{(i+1)} = v_j^{(i)} + \delta v_j^{(i)}$$

$$w_j^{(i+1)} = w_j^{(i)} + \delta w_j^{(i)}$$

$$s_j^{(i+1)} = s_j^{(i)} + \delta s_j^{(i)}$$

$$t_j^{(i+1)} = t_j^{(i)} + \delta t_j^{(i)}$$

4. Block Elimination Technique

In the last step, the results of linearization are solved by using the tridiagonal block matrix elimination technique.

20.3 Results and Discussions

In this paper, the mixed convection of viscoelastic fluid at the lower stagnation point on a magnetic sliced sphere is investigated numerically using Keller box scheme. The fluid is viscoelastic fluid, which has characteristics both of viscous and elastic, with the presence of magnetic field from the sliced magnetic sphere. The objective of the present analysis is to study the temperature profiles and velocity profiles of mixed convection of viscoelastic fluid flow with the variation of magnetic parameter value (*M*), viscoelastic parameter (*K*), mixed convection parameter (λ), Prandtl Number (*Pr*), and the slicing angle of the sphere (θ_s).

20.3.1 The Effects of Magnetic Parameters (*M*)

In Figs. 20.2 and 20.3, the magnetic parameters used are 15, 25, 30, 35, and other parameter values used are $K = 1$, $Pr = 0.7$, $\lambda = 1$, $\theta_s = \frac{\pi}{6} = 30^\circ$. Figure 20.2 shows that when the magnetic parameters increase, then the velocity profile increases.

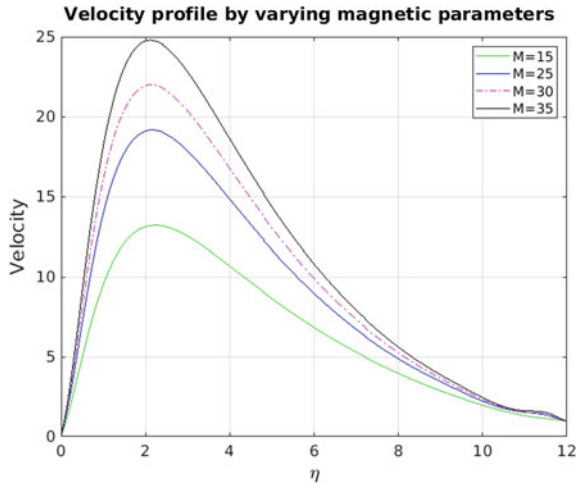


Fig. 20.2 Velocity profile for various magnetic parameter (M) at the lower stagnation point

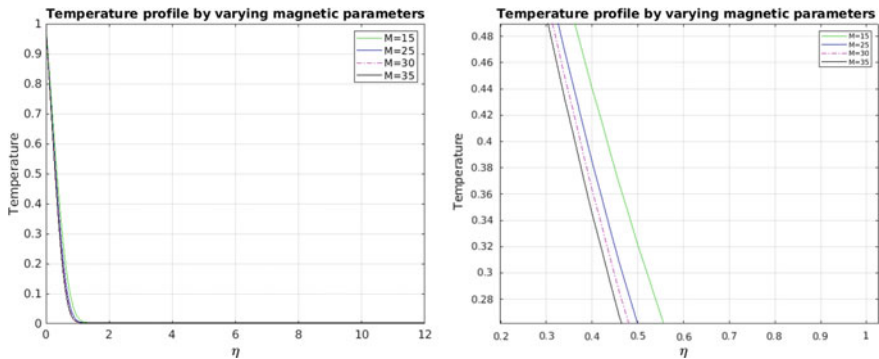


Fig. 20.3 Temperature profile for various magnetic parameter (M) at the lower stagnation point

While, Fig. 20.3 shows that the temperature profile decreases when the magnetic parameter increases. This is because when the magnetic parameter increases then the uniform magnetic field becomes higher. This uniform magnetic field is generated by Lorentz force from the sphere, so the resistance of the forces between fluid particles reduce. And this causes the fluid velocity increases. The uniform magnetic field increases also causes the internal energy of the fluid increases so that the energy used by the fluid flows will be reduced. This effects the temperature decreases.

20.3.2 Influence of Viscoelastic Parameters (K)

Figures 20.4 and 20.5 respectively show the profile of the velocity and temperature of the fluid at the lower stagnation, with the variation of the viscoelastic parameters being 2, 5, 7, 9. This profile is obtained when we give the parameter value of $M = 10$, $Pr = 0.7$, $\lambda = 1$ and $\theta_s = \frac{\pi}{6} = 30^\circ$. From this we obtain numerical calculation results that the higher the viscoelastic parameter, the slower the velocity profile. Moreover, the higher the value of the viscoelastic parameter, the higher the temperature profile. This is because the higher the viscoelastic parameter, the higher the elasticity and viscosity of the fluid. This causes the fluid movement to be slower and will increase the energy requirements for the fluid moving from the bottom to the top and the fluid force against the Lorentz force. Therefore, these results in an increase in the temperature of the fluid.

20.3.3 Influence of Mixed Convection Parameters (λ)

Figures 20.6 and 20.7 respectively show the profile of the velocity and temperature of the fluid at the lower stagnation, with mixed convection parameters being 0.2, 0.4, 0.6, 0.9. This profile is obtained when we give the parameter values of $M = 10$, $Pr = 0.7$, $K = 1$, $\theta_s = \frac{\pi}{6} = 30^\circ$. From this we obtain the results of numerical calculations that the higher the mixed convection parameter, the faster the fluid velocity. Moreover, the higher the value of the mixed convection parameter, the higher the fluid temperature. This is due to the influence of buoyancy, where the buoyant force of the fluid affects

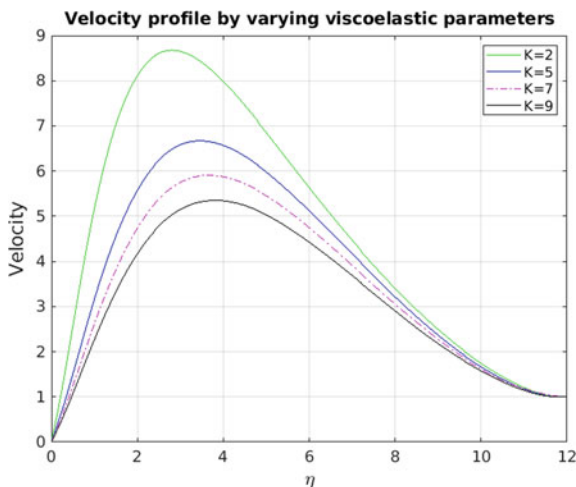


Fig. 20.4 Velocity profile for various viscoelastic parameters (K) at the lower stagnation point

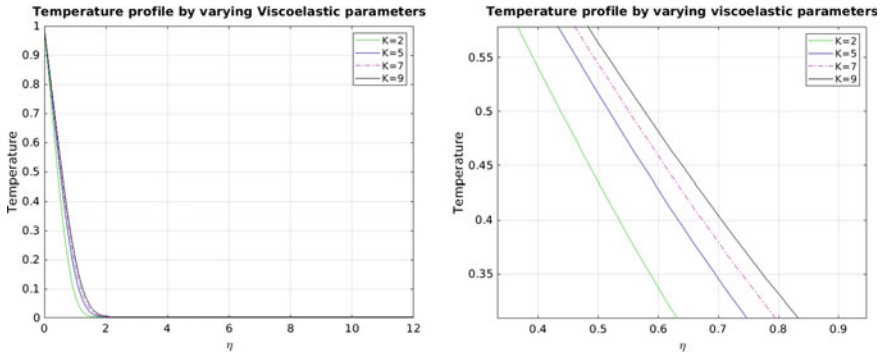


Fig. 20.5 Temperature profile for various viscoelastic parameters (K) at the lower stagnation point

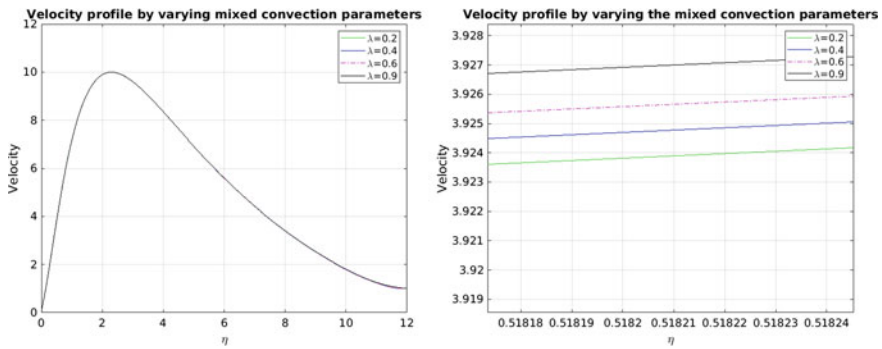


Fig. 20.6 Velocity profile for various mixed convection parameters (λ) at the lower stagnation point

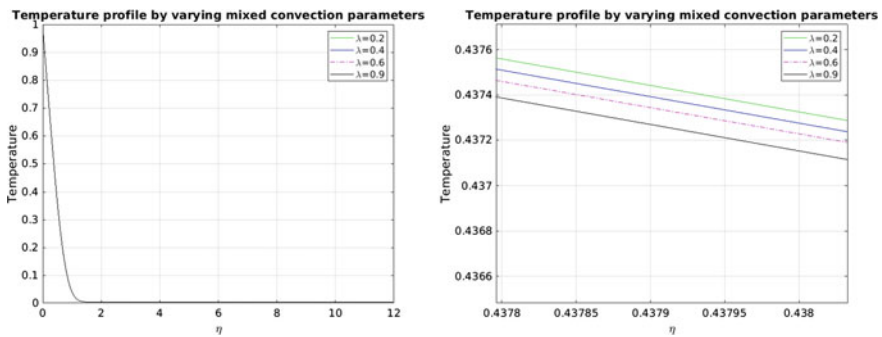


Fig. 20.7 Temperature profile for various mixed convection parameters (λ) at the lower stagnation point

the greater the momentum of the fluid so that the velocity of the fluid flow increases. Furthermore, this causes the need for energy and force to counter the Lorentz force decrease. This causes the temperature near the surface to decrease.

20.3.4 Influence of Prandtl Numbers (Pr)

In Figs. 20.8 and 20.9, respectively, we investigate the effect of changing the Prandtl number. We varied the Prandtl number to 0.6, 1, 7, and 10, while the other parameter values remained constant, namely at $M = 10$, $\lambda = 1$, $K = 1$, $\theta_s = \frac{\pi}{6} = 30^\circ$. Figure 20.8 shows that the larger the Prandtl Number, the slower the speed. Furthermore, Fig. 20.9 shows that the greater the Prandtl Number value, the lower the fluid temperature. This is because the Prandtl number is directly proportional to the density so that if the Prandtl number increases, the density of the fluid also increases. This causes the distance between the fluid particles to be shorter. Moreover, this is the smaller the average momentum of the fluid particles against the Lorentz force of the sphere. This causes the fluid velocity to decrease as the Prandtl number increases. Notice also that the Prandtl number is inversely proportional to the thermal diffusion, which means that the greater the Prandtl number, the less thermal diffusion. Thermal diffusivity is related to the ratio of heat transfer to the energy capacity of molecules so that the distribution of heat between fluids is reduced or in other words heat transfer to the surface is faster than the fluid, causing the temperature decreases.

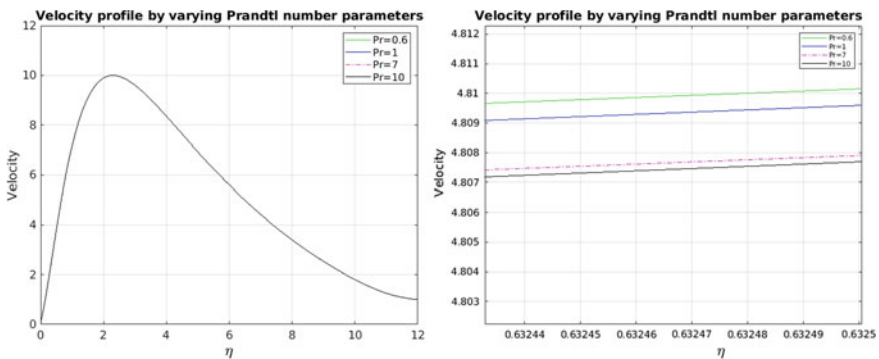


Fig. 20.8 Velocity profile for various Prandtl numbers (Pr) at the lower stagnation point

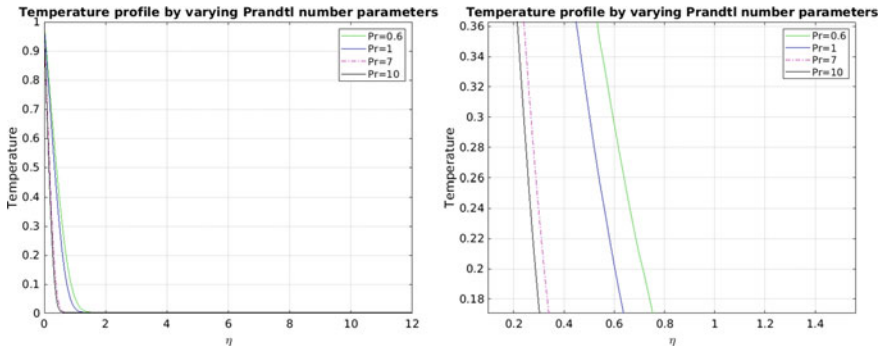


Fig. 20.9 Temperature profile for various Prandtl numbers (Pr) at the lower stagnation point

20.3.5 Influence of Varying the Parameters of Sphere Slice Angle (θ_s)

In Figs. 20.10 and 20.11, respectively, we investigate the effect of changing the angle of the slice of the sphere. We varied the angle of the slice of the sphere to $\frac{\pi}{7}$, $\frac{\pi}{6}$, $\frac{\pi}{5}$ and $\frac{\pi}{4}$, while the other parameter values remained constant, namely at $M = 10$, $\lambda = 1$, $K = 1$, $Pr = 0.7$. Figure 20.10 shows that the greater the angle of the sphere slice, the faster the fluid velocity. Moreover, Fig. 20.11 shows that the greater the angle of the sphere slice, the lower the fluid temperature. This is because the greater the angle of the slice on the sphere causes the position of the stagnation point to be higher or in other words the thickness of the boundary layer is getting bigger. Therefore, this causes the influence of the Lorentz force on the fluid particles decrease so that the fluid velocity increases. Notice also the greater the angle of slice of the sphere,

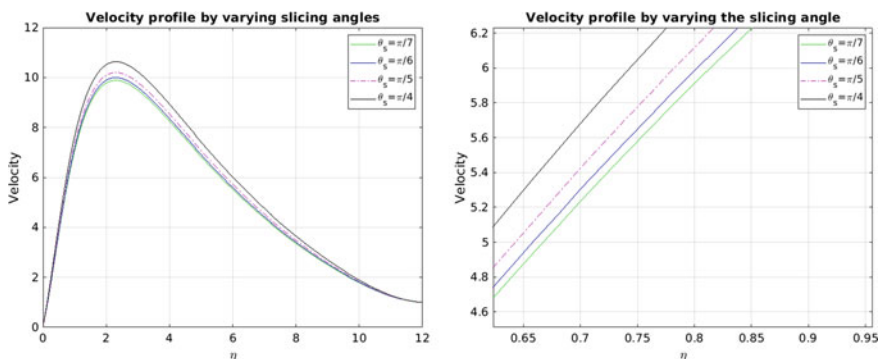


Fig. 20.10 Velocity profile for various sphere slice angles (θ_s) at the lower stagnation point

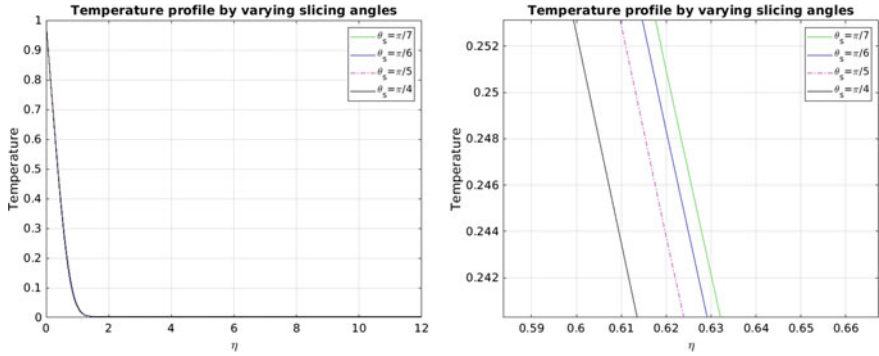


Fig. 20.11 Temperature profile for various sphere slice angles (θ_s) at the lower stagnation point

the wider the front surface of the sphere which causes the heat distribution to the sphere to be faster when compared to the heat distribution to the fluid so that the fluid temperature decreases.

20.4 Conclusions

Based on the discussion and analysis that we have done, it is concluded that the uniform magnetic field produced by Lorentz force and slicing on the sphere act as determining factors for the trend of viscoelastic fluid movement and controlling the cooling rate of the sphere surface.

Acknowledgements Based on the discussion and analysis that we have done, it is concluded that the uniform magnetic field produced by Lorentz force and slicing on the sphere act as determining factors for the trend of viscoelastic fluid movement and controlling the cooling rate of the sphere surface.

References

1. B. Widodo, *Pemodelan Matematika* (ITS Press, Surabaya, 2012)
2. G. Lazarus, *J. Therm. Eng.* **1**(2), 113 (2015)
3. B. Widodo, D. Khalimah, F. Zainal, C. Imron, *Int. J. Far East J. Math. Sci.* (Publishing House-India) (2015)
4. S. Irianto, Ltu operation combination – mhd – fuel cell operation and possible application in Indonesia. Bachelor’s thesis, Institut Teknologi Sepuluh Nopember, Indonesia, 2013
5. P. Davidson, *An Introduction to Magnetohydrodynamics* (Cambridge University Press, Cambridge, 2010)
6. R. Sahaya, Magnetohydrodynamic suspended viscoelastic fluid flow through a flat plate. Bachelor’s thesis, Institut Teknologi Sepuluh Nopember, Indonesia, 2016

7. B. Widodo, A. Sulistyanyngtyas, C. Imron, Effect of viscoelastic fluid flow past an elliptic cylinder. Technical report (Institut Teknologi Sepuluh Nopember, 2015)
8. A. Ishak, R. Nazar, I. Pop, *Magnetohydrodynamics* **45**(1), 103 (2009)
9. A. Kasim, Convective boundary layer flow of viscoelastic fluid. Ph.D. thesis, Universiti Teknologi Malaysia, 2014
10. F. Ali, R. Nazar, N. Arifin, I. Pop, *Appl. Math. Mech.* **32**(4), 409 (2011)
11. K. Jafar, R. Nazar, A. Ishak, I. Pop, *Bound. Value Probl.* **2013**(1), 1 (2013)
12. K. Ali, M. Ashraf, S. Ahmad, K. Batool, *World Appl. Sci. J.* **16**(1), 1638 (2012)
13. K. Ali, M. Ashraf, *World Appl. Sci. J.* **16**(11), 1615 (2012)
14. B. Widodo, P. Wanti, C. Imron, Viscoelastic fluid past a flat plate with the effect of magnetohydrodynamic. Technical report (Institut Teknologi Sepuluh Nopember, 2015)
15. M. Utama, B. Widodo, D. Adzkiya, Unsteady magnetohydrodynamics on mixture convection flowing through a sliced sphere in viscous fluid under the influence of magnetic field. Technical report (Institut Teknologi Sepuluh Nopember, 2017)

Chapter 21

Stability Analysis of Fractional-Order Chemostat Model with Time Delay



Nor Afiqah Mohd Aris and Siti Suhana Jamaian

Abstract The fractional-order chemostat model that considered time delay is studied. Fractional-order differential equations have more benefits for the explanation of memory and the hereditary properties of a system. However, fractional-order differential equations tend to lower the dimensionality of a system. The dimensionality can be infinite-dimensional if the time delay is considered in the differential equation. The stability analysis of the fractional-order chemostat model that considered time delay is studied to examine the effect of time delay on the behaviour of the chemostat system. The numerical simulation was conducted to investigate the fractional-order chemostat model with various values of fractional-order corresponding to different values of time delay. The simulation was performed by using the modified Adams-type predictor–corrector method. The result shows that the stable state transformed into an unstable state or a limit cycle at an appropriate time delay value. As the fractional-order value decreased, a higher value of time delay has to be considered. Therefore, the suitable time delay value can be appropriately selected to ensure that the chemostat system’s dynamic behaviour is constantly unstable, which is appropriate for cell mass growth.

21.1 Introduction

In the past few decades, many researchers have studied improving cell mass production in chemical reactors. The chemostat model can understand the mechanism of cell mass growth in a chemostat [1]. A chemostat is a reactor with stirred and continuous inflow and outflow to provide efficient mixing. The chemostat design has an inflow tube for the substrate to flow into the chemostat and an outflow tube for harvesting the mixture of cell mass and substrate from the chemostat while the volume is kept

N. A. Mohd Aris (✉) · S. S. Jamaian
Department of Mathematics and Statistics, Faculty of Applied Sciences and Technology,
Universiti Tun Hussein Onn Malaysia, Pagoh Education Hub, 84600 Pagoh, Johor, Malaysia

S. S. Jamaian
e-mail: suhana@uthm.edu.my

constant. The substrate flows sufficiently rapid to ensure that the culture of the chemostat model keeps continuously growing. The process of the continuous growth of cell mass occurs in the reactor with the consumption of the substrate at a specific growth rate. The chemostat model was first introduced by Novick and Szilard [2]. However, the chemostat model that was wholly derived and only considered a single substrate and a single microbial was proposed by Herbert, Elsworth and Telling [3].

The microbial kinetics model explains the relationship between cell mass and the substrate in the chemostat model. Many microbial kinetics models exist, such as the Monod, Tessier, Moser, Contois and Andrews models [4–7]. The yield coefficient is a function of the substrate concentration, which can express the cell mass formation. The yield coefficient can be demonstrated as the cell mass formed per unit mass of the substrate consumed [8]. Some studies also considered the variable yield coefficient in investigating the cell mass growth in a chemostat [9, 10]. A model with a variable yield coefficient is suitable for developing a continuous culture [11].

Most mathematical models of biological systems are based on ordinary differential equations. However, the behaviour of biological systems mostly has memory effects. Yet, ordinary differential equations consistently disregard such effects. Therefore, fractional-order differential equations are used to describe the behaviour of systems. Furthermore, the fractional-order differential equations also have hereditary properties and good memory [12, 13]. The capability of fractional-order differential equations to describe non-linear systems in various fields of study is also advantageous. However, the practical dimension in fractional-order differential equations for which the system remains unstable is also one of the crucial problems in this field [14]. The fractional-order differential equations tend to lower the dimensionality of a system [15]. Meanwhile, the dimensionality can be infinite-dimensional if the time delay is considered in the differential equations.

Delay differential equations (DDE) are widely applied in the mathematical modelling of biological systems such as immunology, neural networks, epidemiology, physiology and population dynamics [16–21]. There are some durations of hidden processes, which can be related to time delays. Derivatives and unknown states are evaluated at the exact moment in ordinary differential equations. However, derivatives and unknown states are estimated at a specific time, depending on the memory, in delay differential equations [22]. The function values at the previous time affect the derivatives of the function values at present [23]. Also, Kuang [23] stated that function values depend on their history, not only on their present state. Initial conditions and boundary conditions are not sufficient to predict the behaviour of a system in real-life phenomena. Hence, some knowledge of history or earlier behaviour is necessary to deal with such complexities.

Delay differential equations have constantly appeared in many studies of the chemostat system [24–28]. The time delays in substrate recycling in the chemostat ecosystem always exist and are also influenced by the changes in temperature. The environment in the chemostat is the crucial factor that affects cell mass growth. The time delay of the growth appears when the environment changes. The time delay has biological values since it can alter the solutions of the chemostat system's dynamics, and the transient oscillation can be displayed in numerical simulations [29]. A delay

can express the information from the earlier state and provide a history of the system over the delay interval $[-\tau, 0]$ as the initial condition where τ is time delay [14, 30]. These cause delay systems to be infinite dimensions in nature. Sometimes, internal and external uncertainties in the application, such as time delay, cannot be avoided and may lead to instability [31]. Delay can be recognised everywhere and is often encountered in many practical systems such as biology, economics and automatic control [32]. This study intends to deepen the study of the integer-order chemostat model with fractional-order theory and consider the time delay due to its broad applicability. In the next section, methods of this research are discussed in more detail, where Caputo derivatives and Adams-type predictor–corrector method are the main focus. Then, the stability analysis of the fractional-order chemostat model with time delay is discussed.

21.2 Methodology

21.2.1 Caputo Derivatives

The fractional-order differential equations can be defined as the rationalisation of classical differential equations involving the differentiation and integration of non-linear order [33]. A fractional-order differential equation has several definition operators to generalise the notion of differentiation and integration, such as Riemann–Liouville definitions, Grunwald–Letnikov definitions and many more. Caputo definitions are more renowned in the related theories and mathematical analysis of fractional-order differential equations than the others [34, 35]. Here we adopt the Caputo fractional derivative, which can be defined as

$$D_{a,t}^{\alpha} f(t) = \int_a^t (\tau - t)^{n-\alpha-1} f^{(n)}(\tau) d\tau, \quad n - 1 \leq \alpha < n \quad (21.1)$$

where n is an integer, Γ denotes the function of gamma and $n - 1 < \alpha < n$.

21.2.2 Adams-Type Predictor-Corrector Method

This research adopted the Adams-type predictor–corrector method that was used by [36] and adjusted by [37] with time delay, which is also known as the modified Adams-type predictor–corrector method [15, 38]. The predictor formula can be described as

$$y_h^P(t_{n+1}) = \sum_{k=0}^{\lceil \alpha \rceil - 1} \frac{t_{n+1}^k}{k!} y_0^{(k)} + \frac{1}{\Gamma(\alpha)} \sum_{j=0}^n b_{j,n+1} f(t_j, y_h(t_j)) \tag{21.2}$$

and corrector formula can be determined by

$$y_h(t_{n+1}) = \sum_{k=0}^{\lceil \alpha \rceil - 1} \frac{t_{n+1}^k}{k!} + \frac{h^\alpha}{\Gamma(\alpha + 2)} f(t_{n+1}, y_h^P(t_{n+1})) + \frac{h^\alpha}{\Gamma(\alpha + 2)} \sum_{j=0}^n a_{j,n+1} f(t_j, y_h(t_j)), \tag{21.3}$$

where h is step size, k is iteration, t is time, α is an order of fractional derivatives, Γ is gamma, τ is the time delay, $a_{j,n+1}$ is corrector weight, and $b_{j,n+1}$ is predictor weight. The procedure of the predictor–corrector method for solving fractional-order differential equations with a time delay can be explained as follows.

- (i) Insert the initial condition of the system.
- (ii) Define the step size, h .
- (iii) Calculate the predictor step, $y_h^P(t_{n+1})$ as in Eq. (21.2).
- (iv) Evaluate $f(t_{n+1}, y_h^P(t_{n+1}))$ by implementing the predictor value from step (iii).
- (v) Calculate the corrector step, $y_h(t_{n+1})$ as in Eq. (21.3).
- (vi) Evaluate $f(t_j, y_h(t_j))$ by implementing the predictor value from step (v).
- (vii) Estimate the error tolerance.

The process predicts and corrects the values repeatedly until the corrected values become converged [39].

21.3 Results and Discussion

An integer-order chemostat model from [1] is studied. The chemostat model can be described as

$$\frac{dS}{dt} = QS_0 - QS - \frac{N(S)X}{Y}, \tag{21.4}$$

$$\frac{dX}{dt} = QX_0 - QX + N(S)X, \tag{21.5}$$

where X is cell mass concentration, X_0 is initial cell mass concentration, S is the substrate concentration, S_0 is initial substrate concentration, Q is dilution rate, $N(S)$

is the specific growth rate and Y is yield coefficient. The Monod growth model depends on the substrate concentration and can be written as,

$$N(S) = \frac{\mu S}{K + S}. \quad (21.6)$$

where μ is the maximum specific growth rate and K is Monod constant. The variable yield coefficient can be written as

$$Y(S) = \gamma + \beta S. \quad (21.7)$$

where γ and β is constant in yield coefficient. In this study, the sterile feed case was assumed, where the initial concentration of the cell mass is equal to zero, or $X_0 = 0$. The chemostat model that considers the Monod growth model and variable yield coefficient can be written as

$$\frac{dS(t)}{dt} = Q(S_0 - S(t)) - \frac{\mu S(t)X(t)}{(k + S(t))(\gamma + \beta S(t))}. \quad (21.8)$$

$$\frac{dX(t)}{dt} = Q(-X(t)) + \frac{\mu S(t - \tau)X(t - \tau)}{k + S(t - \tau)} \quad (21.9)$$

where τ is a time delay. The model incorporates delay to represent the delay in the cell mass concentration, as shown in Eq. (21.9). The integer-order chemostat system of Eqs. (21.8) and (21.9) is extended to the fractional-order differential equation with time delay,

$$\frac{d^\alpha S(t)}{dt^\alpha} = Q(S_0 - S(t)) - \frac{\mu S(t)X(t)}{(k + S(t))(\gamma + \beta S(t))}, \quad (21.10)$$

$$\frac{d^\alpha X(t)}{dt^\alpha} = Q(-X(t)) + \frac{\mu S(t - \tau)X(t - \tau)}{k + S(t - \tau)}. \quad (21.11)$$

The chemostat systems with time delay are extended to FDEs. This research intends to deepen the study of the integer-order chemostat model with fractional-order theory due to its broad applicability. The model incorporated with delay only occurs in the cell mass concentration, as shown in Eq. (21.11). The parameter values of the fractional-order chemostat model are provided in Table 21.1.

The simulations of a fractional-order differential equation that considers time delay are challenging to analyse analytically because of the infinite dimensionality of the DDE [40]. Numerical solutions play an essential role to study the effect of time delay on the fractional-order chemostat system.

The numerical solution of a fractional-order chemostat model with time delay was simulated, and the behaviour of the system was investigated. Figure 21.1 illustrates

Table 21.1 Parameters of chemostat model

Parameters	Values	References
k	1.75	[12]
Q	0.02	[12]
μ	0.3	[1]
γ	0.01	[1]
β	5.25	[1]
S_0	1	[1]

the effect of time delay on the behaviour of the chemostat system when $\alpha = 1$. The phase portrait diagrams of cell mass concentration versus substrate concentration were plotted for different time delays, which were $\tau = 1$ to $\tau = 12$. The phase portrait diagrams $\tau > 12$ are not be displayed since the result show the same behaviour as $\tau = 12$.

From Fig. 21.1a–l, the changes in the running state for $\alpha = 1$ are shown. It is observed that the fractional-order chemostat system is in a stable state when $0 < \tau \leq 5$, as shown in Fig. 21.1a, e. When the value of time delay increased over the limit, which is $\tau = 6$, the dynamic behaviour of the chemostat system led to a limit cycle. The limit cycle then shifts to an unstable state when $\tau \geq 7$ and being unchanged. These results demonstrate that the stable state can be transformed into an unstable state or a limit cycle with a suitable choice of the time delay parameter. From the publications of Sterman and Ydstie [41] and Douglas and Rippin [42], the periodic operation or unstable state of chemical reactors to improve reactor performance has been a common research subject [43]. The “limit cycle” is also known as “natural oscillation” in chemical processes. Natural oscillation is mean that the process parameters are intended to be selected so that a steady input of reactant into the reactor will produce self-sustained oscillations in its output [1]. The limit cycle and unstable state are suitable for cell mass production when the reactor’s performance is improved. It can be considered that the time delay remains at $\tau \geq 6$ so that the chemostat system will always be at an unstable state or limit cycle.

Figure 21.2 depicts the effect of time delay on the behaviour of the chemostat system when $\alpha = 0.95$. The phase portrait diagrams of cell mass concentration against substrate concentration are visualised for different time delay values, which are $\tau = 6$ to $\tau = 17$. The phase portrait diagrams for $\tau < 6$ and $\tau \geq 17$ are not be displayed since the result shows the same behaviour as $\tau = 6$ and $\tau = 17$, respectively.

The change in the running state of the fractional-order chemostat system for $\alpha = 0.95$ is shown in Fig. 21.2. The fractional-order chemostat system showed different running states for $\alpha = 1$ and $\alpha = 0.95$ with different values of time delay. The fractional-order chemostat system was a limit cycle when $\tau = 6$ as shown in Fig. 21.1f. However, when $\alpha = 0.95$, it is revealed, the fractional-order chemostat system is still at the stable state when $\tau = 6$ which is still stable in range $0 < \tau \leq 11$ and changed to the limit cycle when $\tau = 12$. Then, the system switches to an unstable state when $\tau \geq 13$. It is reasonable to conclude that the fractional-order chemostat

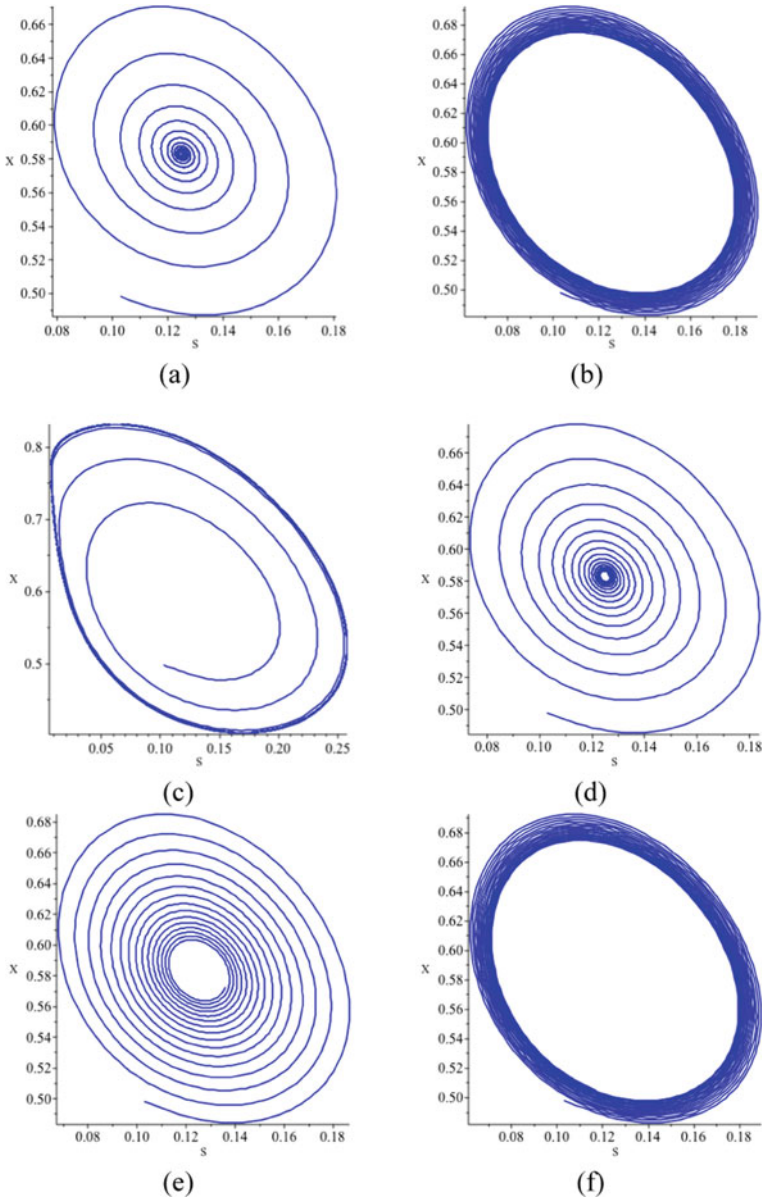
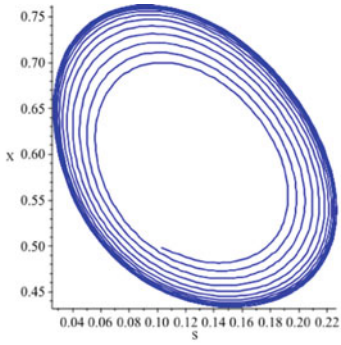
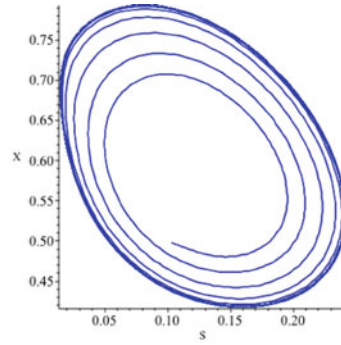


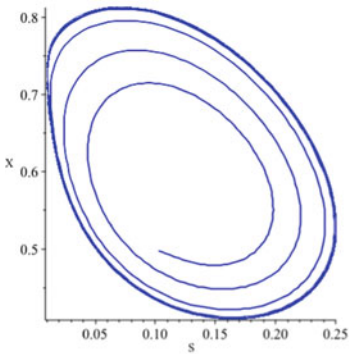
Fig. 21.1 Phase portrait of chemostat system with $\alpha = 1$ and **a** $\tau = 1$, **b** $\tau = 2$, **c** $\tau = 3$, **d** $\tau = 4$, **e** $\tau = 5$, **f** $\tau = 6$, **g** $\tau = 7$, **h** $\tau = 8$, **i** $\tau = 9$, **j** $\tau = 10$, **k** $\tau = 11$, **l** $\tau = 12$



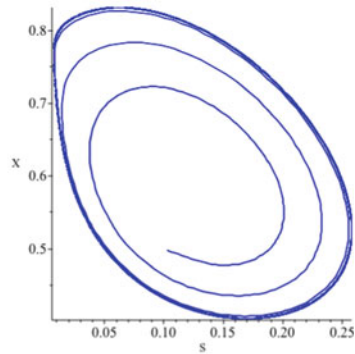
(g)



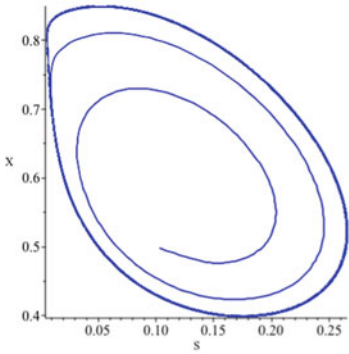
(h)



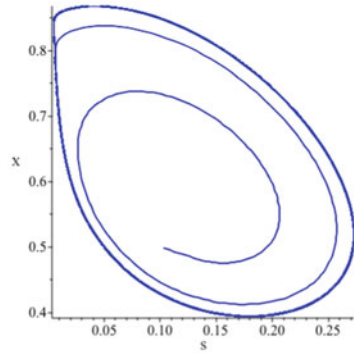
(i)



(j)



(k)



(l)

Fig. 21.1 (continued)

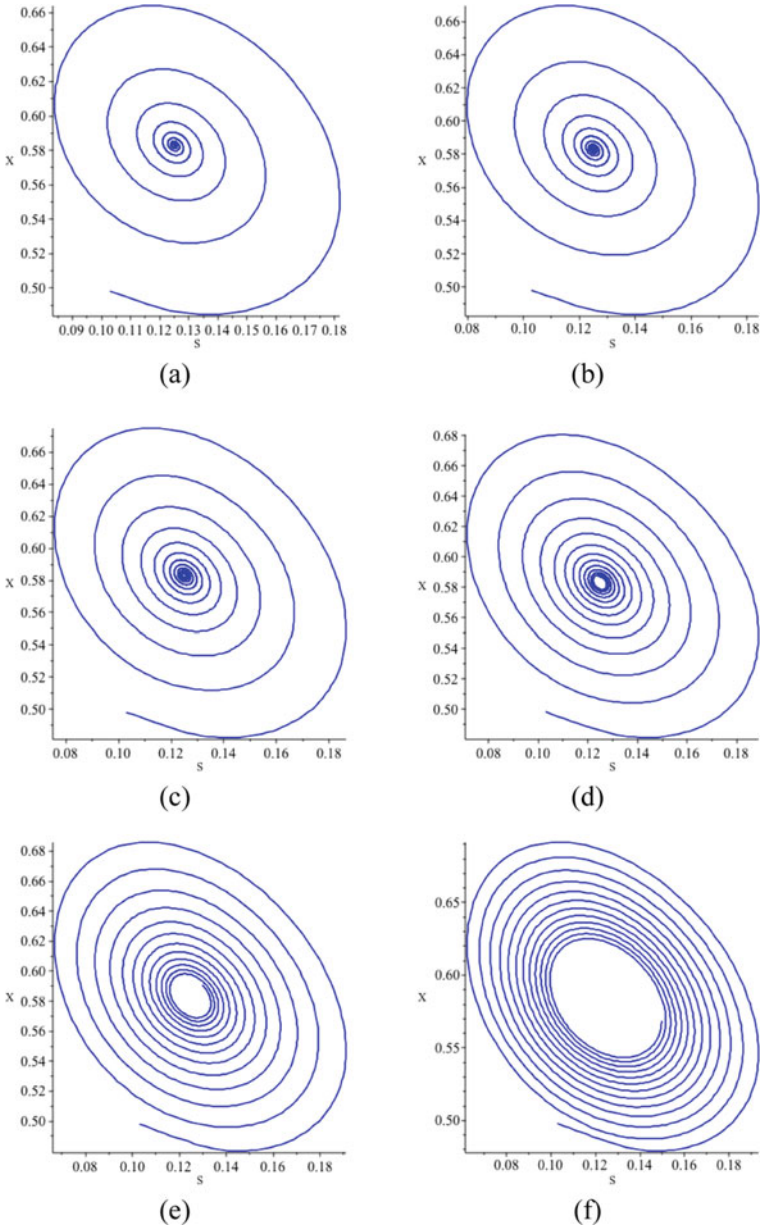
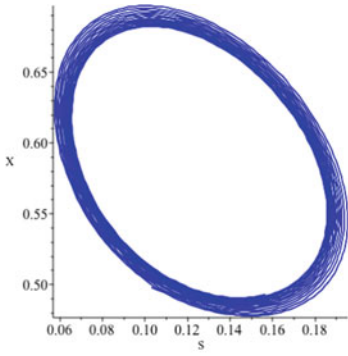
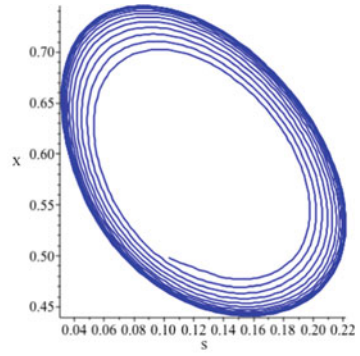


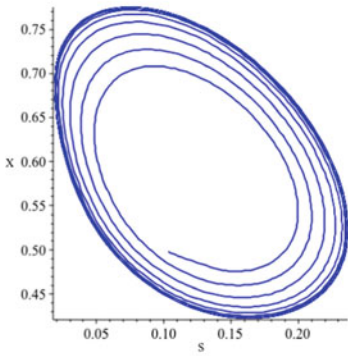
Fig. 21.2 Phase portrait of chemostat system with $\alpha = 0.95$ and **a** $\tau = 6$, **b** $\tau = 7$, **c** $\tau = 8$, **d** $\tau = 9$, **e** $\tau = 10$, **f** $\tau = 11$, **g** $\tau = 12$, **h** $\tau = 13$, **i** $\tau = 14$, **j** $\tau = 15$, **k** $\tau = 16$, **l** $\tau = 17$



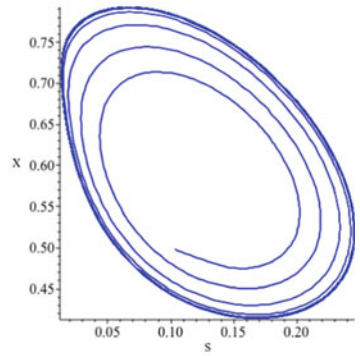
(g)



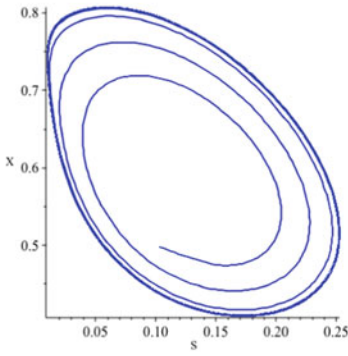
(h)



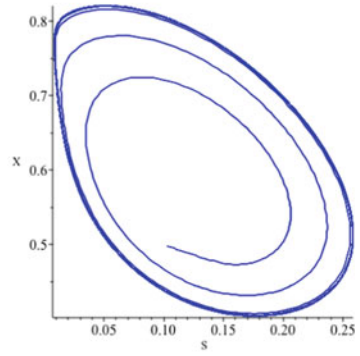
(i)



(j)



(k)



(l)

Fig. 21.2 (continued)

system needed a higher value of time delay when the order of the fractional-order system became smaller. In a real-life application, it can be considered that the time delay remains at $\tau \geq 12$ for $\alpha = 0.95$ so that the fractional-order of the chemostat system will always be at the unstable state.

Figure 21.3 depicts the effect of time delay on the behaviour of the chemostat system when $\alpha = 0.9$. The phase portrait diagrams of cell mass concentration versus substrate concentration are plotted for varied time delay, which is $\tau = 13$ to $\tau = 24$. The phase portrait diagrams for $\tau < 13$ and $\tau > 24$ are not displayed since the result shows the same behaviour as $\tau = 13$ and $\tau = 24$, respectively.

From Fig. 21.3, the change in the running state for $\alpha = 0.9$ is shown. It is observed that the fractional-order chemostat system is at a stable state when $0 < \tau \leq 20$. Meanwhile, the dynamic behaviour of the fractional-order chemostat system changed to the limit cycle when the value of the time delay is increased to $\tau = 21$. Then, the system switches to an unstable state when $\tau \geq 22$. In a real-life application, it can be considered that the time delay remains at $\tau \leq 21$ for $\alpha = 0.9$ so that the dynamic behaviour of the chemostat system will always be unstable and hence is suitable for the production of cell mass. A summary of the observation for different values of α and the time delay is presented in Table 21.2.

21.4 Conclusion

The stability analysis of the fractional-order chemostat model with various time delay values has been conducted with different values of the order of the fractional-order system, which were $\alpha = 1$, $\alpha = 0.95$ and $\alpha = 0.9$. As the evidence from the phase portrait plots, it has been demonstrated that a stable state could be transformed into an unstable state or a limit cycle with a suitable choice of time delay value. Based on the facts in [43], the limit cycle or unstable state is ideal for producing the cell mass. Thus, the stability analysis is required to determine the suitable time delay value for the limit cycle and unstable state. It also can be concluded that when the order value of the fractional chemostat system became smaller, a higher value of the time delay has to be considered.

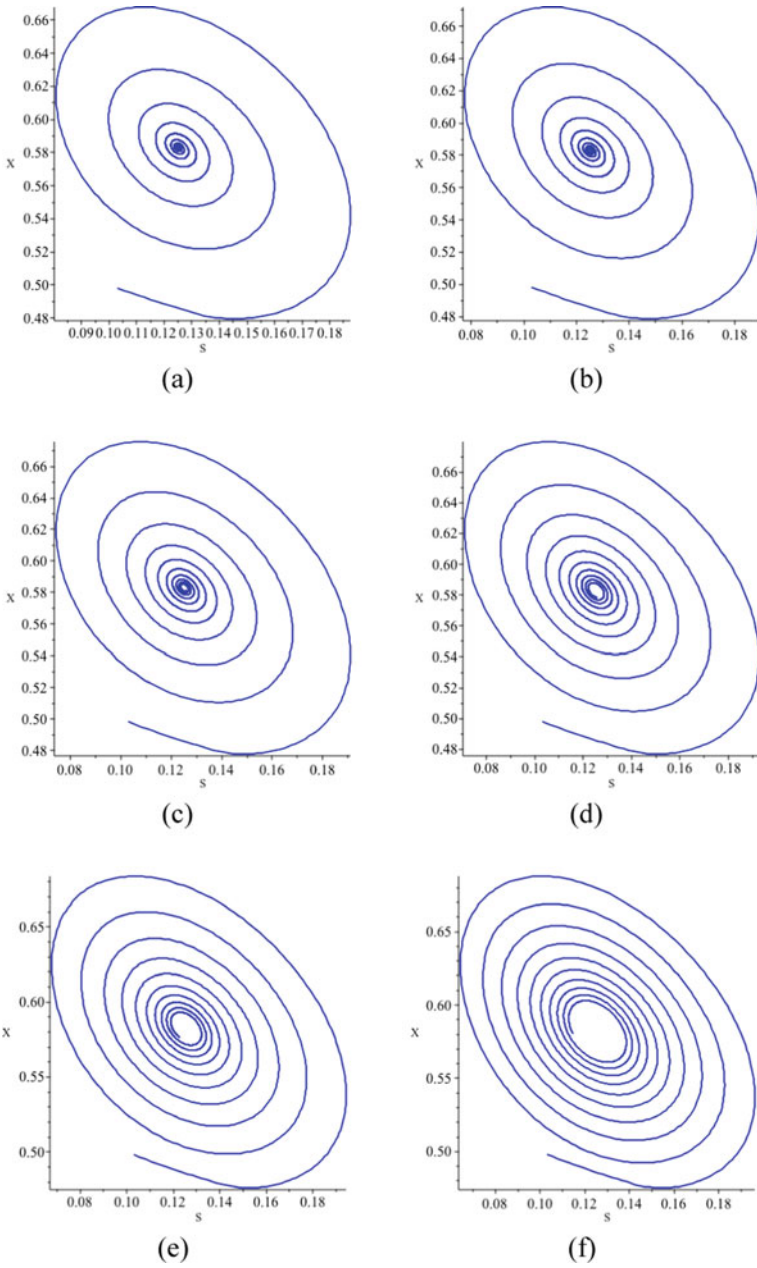


Fig. 21.3 Phase portrait of chemostat system with $\alpha = 0.9$ and **a** $\tau = 13$, **b** $\tau = 14$, **c** $\tau = 15$, **d** $\tau = 16$, **e** $\tau = 17$, **f** $\tau = 18$, **g** $\tau = 19$, **h** $\tau = 20$, **i** $\tau = 21$, **j** $\tau = 22$, **k** $\tau = 23$, **l** $\tau = 24$

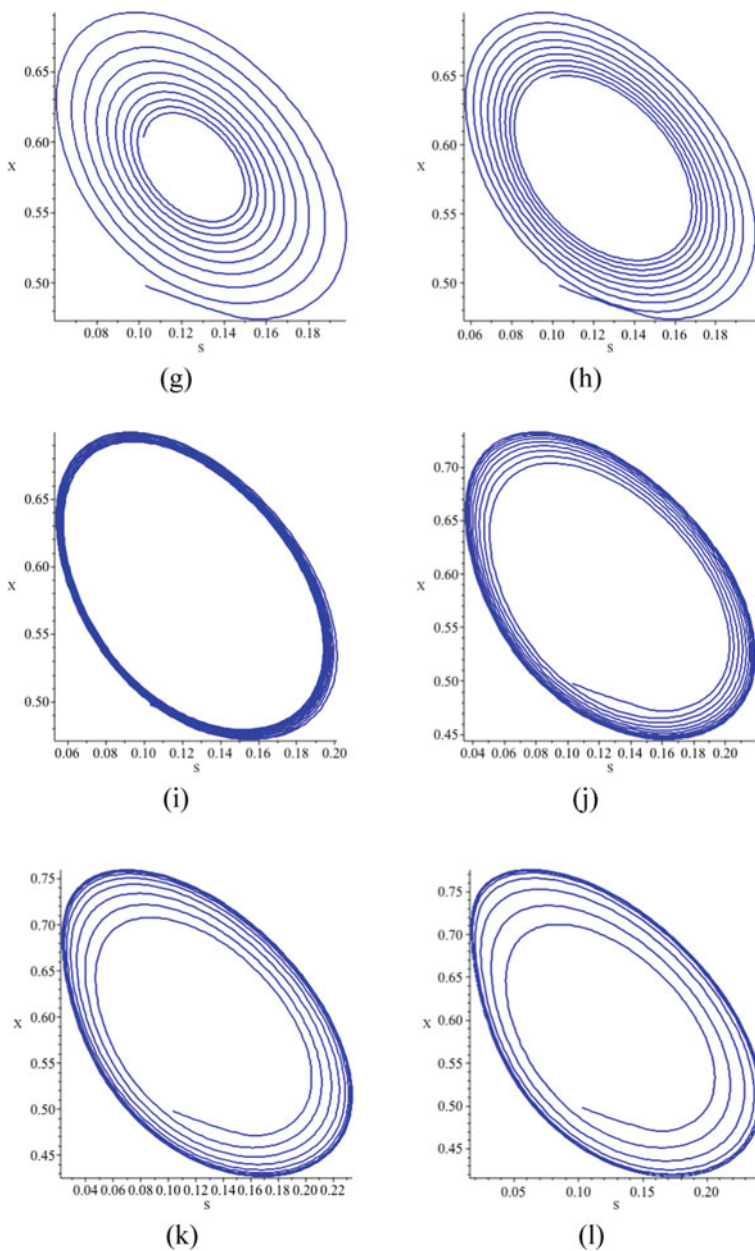


Fig. 21.3 (continued)

Table 21.2 Summary of observation

α	τ	Observation	Figure
1	$0 < \tau \leq 5$	Stable	21.1a–e
	$\tau = 6$	Limit cycle	21.1f
	$\tau \geq 7$	Unstable	21.1g–l
0.95	$0 < \tau \leq 11$	Stable	21.2a–f
	$\tau = 12$	Limit cycle	21.2g
	$\tau \geq 13$	Unstable	21.2h–l
0.9	$0 < \tau \leq 20$	Stable	21.3a–h
	$\tau = 21$	Limit cycle	21.3i
	$\tau \geq 22$	Unstable	21.3j–l

Acknowledgements We would like to express our gratitude sincere for the financial support by the Universiti Tun Hussein Onn Malaysia through the grant H426.

References

1. M.I. Nelson, H.S. Sidhu, Analysis of a chemostat model with variable yield coefficient. *J. Math. Chem.* **38**(4), 605–615 (2005)
2. A. Novick, L. Szilard, Experiments with the chemostat on spontaneous mutations of bacteria. *Proc. Nat. Acad. Sci. U.S.A.* **36**(12), 708 (1950)
3. D. Herbert, R. Elsworth, R.C. Telling, The continuous culture of bacteria; a theoretical and experimental study. *Microbiology* **14**(3), 601–622 (1956)
4. J. Monod, The growth of bacterial cultures. *Annu. Rev. Microbiol.* **3**(1), 371–394 (1949)
5. H. Moser, *The Dynamics of Bacterial Populations Maintained in the Chemostat*, 1st edn. (Carnegie Institution of Washington, United States, 1958)
6. D.E. Contois, Kinetics of bacterial growth: relationship between population density and specific growth rate of continuous cultures. *Microbiology* **21**(1), 40–50 (1959)
7. J.F. Andrews, A mathematical model for the continuous culture of microorganisms utilising inhibitory substrates. *Biotechnol. Bioeng.* **10**(6), 707–723 (1968)
8. J. Hong, Yield coefficients for cell mass and product formation. *Biotechnol. Bioeng.* **33**(4), 506–507 (1989)
9. I.H. Lee, A.G. Fredrickson, H.M. Tsuchiya, Damped oscillations in continuous culture of *Lactobacillus plantarum*. *J. Gen. Microbiol.* **93**(2), 204–208 (1976)
10. A.G. Dorofeev, M.V. Glagolev, T.F. Bondarenko, N.S. Panikov, Observation and explanation of the unusual growth-kinetics of arthrobacter-globiformis. *Microbiology* **61**(1), 24–33 (1992)
11. S. Suzuki, K. Shimizu, M. Matsubara, On the parameter-space classification of the dynamic behavior of a continuous microbial flow reactor. *Chem. Eng. Commun.* **33**(5–6), 325–335 (1985)
12. M. Zeinadini, M. Namjoo, A numerical method for discrete fractional-order chemostat model derived from nonstandard numerical scheme. *Bull. Iran. Math. Soc.* **43**(5), 1165–1182 (2017)
13. M. Zeinadini, M. Namjoo, Approximation of fractional-order chemostat model with nonstandard finite difference scheme. *Hacettepe J. Math. Stat.* **46**(3), 469–482 (2017)
14. S. Bhalekar, V. Daftardar-Gejji, Fractional ordered Liu system with time-delay. *Commun. Non-linear Sci. Numer. Simul.* **15**(8), 2178–2191 (2010)

15. V. Daftardar-Gejji, S. Bhalekar, P. Gade, Dynamics of fractional-ordered chen system with delay. *Pramana* **79**(1), 61–69 (2012)
16. A. Isah, C. Phang, Operational matrix based on Genocchi polynomials for solution of delay differential equations. *Ain Shams Eng. J.* **9**(4), 2123–2128 (2018)
17. G.A. Bocharov, F.A. Rihan, Numerical modelling in biosciences using delay differential equations. *J. Comput. Appl. Math.* **125**(1–2), 183–199 (2000)
18. F.A. Rihan, D.H. Abdelrahman, F. Al-Maskari, F. Ibrahim, M.A. Abdeen, Delay differential model for tumour-immune response with chemoimmunotherapy and optimal control. *Comput. Math. Methods Med* (2014)
19. C.T.H. Baker, G.A. Bocharov, C.A.H. Paul, F.A. Rihan, Modelling and analysis of time-lags in some basic patterns of cell proliferation. *J. Math. Biol.* **37**(4), 341–371 (1998)
20. S. Lakshmanan, F.A. Rihan, R. Rakkiyappan, J.H. Park, Stability analysis of the differential genetic regulatory networks model with time-varying delays and Markovian jumping parameters. *Non-linear Anal: Hybrid Syst.* **14**, 1–15 (2014)
21. R. Rakkiyappan, G. Velmurugan, F.A. Rihan, S. Lakshmanan, Stability analysis of memristor-based complex-valued recurrent neural networks with time delays. *Complexity* **21**(4), 14–39 (2016)
22. F.A. Rihan, C. Tunc, S.H. Saker, S. Lakshmanan, R. Rakkiyappan, Applications of delay differential equations in biological systems. *Complexity* (2018)
23. Y. Kuang, *Delay differential equations: with applications in population dynamics* (Academic press, United States of America, 1993)
24. X. Zhang, S. Sun, Dynamical analysis of a stochastic delayed two-species competition chemostat model. *Bull. Malays. Math. Sci. Soc.* **43**, 3725–3755 (2020)
25. S. Wang, S. He, A. Yousefpour, H. Jahanshahi, R. Repnik, M. Perc, Chaos and complexity in a fractional-order financial system with time delays. *Chaos, Solitons Fractals* **131** (2020)
26. P. Amster, G. Robledo, D. Sepúlveda, Existence of ω -periodic solutions for a delayed chemostat with periodic inputs. *Non-linear Anal.: Real World Appl.* **55** (2020)
27. J. Ileri, G. Pokhariyal, S. Moindi, Hopf bifurcation analysis for a two species periodic chemostat model with discrete delays. *J. Adv. Math. Comput. Sci.* **35**(3), 93–105 (2020)
28. Y. Li, Hopf bifurcation of the chemostat with delay and simplified Holling type-iv response function. *EDP Sci.* **309** (2020)
29. T. Zhang, X. Ji, B. Li, Periodic oscillations in a chemostat model with two discrete delays. *Discr. Dyn. Nat. Soc.* (2015)
30. A. Jhinga, V. Daftardar-Gejji, A new numerical method for solving fractional delay differential equations. *Comput. Appl. Math.* **38**(4), 166 (2019)
31. X. Cui, Y. Yu, H. Wang, W. Hu, Dynamical analysis of memristor-based fractional-order neural networks with time delay. *Mod. Phys. Lett. B* **30**(18) (2016)
32. Z. Wang, A numerical method for delayed fractional-order differential equations. *J. Appl. Math.* (2013)
33. M.P. Lazarević, M.R. Rapaić, T.B. Šekara, V. Mladenov, N. Mastorakis, Introduction to fractional calculus with brief historical background, in *Advanced Topics on Applications of Fractional Calculus on Control Problems, System Stability and Modeling* (World Scientific and Engineering Academy and Society Press, Belgrade, 2014), pp. 3–16
34. H. Mao, D. Zhao, F. Jing, H. Liu, X. Wei, Propagation characteristics of the kurtosis parameters of flat-topped beams passing through fractional Fourier transformation systems with a spherically aberrated lens. *J. Opt. A: Pure Appl. Opt.* **6**(6), 640 (2004)
35. C. Li, G. Chen, Chaos in the fractional order Chen system and its control. *Chaos, Solitons Fractals* **22**(3), 549–554 (2004)
36. K. Diethelm, An algorithm for the numerical solution of differential equations of fractional order. *Electron. Trans. Numer. Anal.* **5**(1), 1–6 (1997)
37. K. Diethelm, N.J. Ford, A.D. Freed, A predictor-corrector approach for the numerical solution of fractional differential equations. *Non-linear Dyn.* **29**(1–4), 3–22 (2002)
38. S. Bhalekar, V. Daftardar-Gejji, A predictor-corrector scheme for solving non-linear delay differential equations of fractional order. *J. Fractional Calc. Appl.* **1**(5), 1–9 (2011)

39. Y.T. Toh, C. Phang, J.R. Loh, New predictor-corrector scheme for solving non-linear differential equations with Caputo-Fabrizio operator. *Math. Methods Appl. Sci.* **42**(1), 175–185 (2019)
40. J.K. Hale, S.M.V. Lunel, *Introduction to Functional Differential Equations* (Springer Science & Business Media, United State of America, 1993)
41. L.E. Stermann, B.E. Ydstie, Periodic forcing of the CSTR: an application of the generalised II-criterion. *AIChE J.* **37**(7), 986–996 (1991)
42. J.M. Douglas, D.W.T. Rippin, Unsteady state process operation. *Chem. Eng. Sci.* **21**(4), 305–315 (1996)
43. R.Y.K. Yang, J. Su, Improvement of chemostat performance via non-linear oscillations. *Bioprocess. Eng.* **9**(2–3), 97–102 (1993)

Chapter 22

Free Convection Nanofluid Flow Near a Three-Dimensional Stagnation Point Induced by g-Jitter: Constant Heat Flux



Mohamad Hidayad Ahmad Kamal, Anati Ali, Noraihan Afiqah Rawi, and Sharidan Shafie

Abstract A theoretical study on free convection boundary layer nanofluid flow near a three-dimensional stagnation point body was conducted under microgravity environment subjected to heat flux boundary condition. g-Jitter effect occurs under microgravity environment caused a fluctuation gravitational field on the fluid system consist of copper nanoparticle and water base fluid. The flow problem was governed into a system of partial differential equation and solved numerically using implicit finite different scheme. The result was presented graphically in term of skin friction coefficient and Nusselt number. The results show that curvature ratio parameter that represent the geometrical shape of boundary body produced different types of stagnation point flow cases. In addition, the increases of amplitude modulation parameter values increase the fluctuation amplitude of each skin friction coefficients and Nusselt number while contradict behavior were noticed as frequency of oscillation parameter increased. Besides that, increases of nanoparticle volume fraction increase both skin friction coefficients and Nusselt number.

22.1 Introduction

In a way to understand fluid behavior and its characteristic, theoretical study is an important sector that need to be considered since it provides a reliable result that can be used as pioneer platform in conducting the study. As Prandtl [1] introduced boundary layer flow that the problem was reduced from mathematical model described by Navier and Stokes, there are a lot of study conducted in understanding

M. H. A. Kamal · A. Ali · N. A. Rawi · S. Shafie (✉)
Department of Mathematical Sciences, Faculty of Science, Universiti Teknologi Malaysia (UTM),
81310 Johor Bahru, Johor Darul Takzim, Malaysia
e-mail: sharidan@utm.my

A. Ali
e-mail: anati@utm.my

N. A. Rawi
e-mail: noraihanafiqah@utm.my

these properties better [2–6]. By specifying the fluid problem into the boundary layer flow, most of the times the fluid problems are subjected to common boundary condition such as constant wall temperature (CWT), constant heat flux (CHF), Newtonian heating (NH) and convective boundary condition (CBC) [7]. The CHF boundary condition is described in many practical and experimental circumstances, where the convective flows is uniformly generated adjacent to boundary body. In experiments, the arrangement for the study of flow over a CHF surface is much simpler than that which employs heaters to obtain a given temperature variation at the surface. The study on CHF has been explored widely since it was introduced with different types of fluids, geometry and fluid effects decades ago. Merkin and Pop [8] studied on free convection boundary layer flow on a horizontal circular cylinder with constant heat flux where the study was focused on comparing the accuracy of Blasius expansion and Gortler-type expansion. Then, Alazmi and Vafai [9] investigated constant heat flux in porous medium under local thermal non-equilibrium (LTNE). Recently, Ahmad et al. [10] investigated mixed convection flow over a horizontal circular cylinder in viscoelastic fluid with CHF in presence of radiation effect. Alsabery et al. [11] numerically studied the magnetohydrodynamics (MHD) natural convection in oblique porous cavity with CHF. Finite element method was used to solve the non-dimensional governing equations.

Introduction of nanofluid usage in boundary layer application shown an impressive results where better fluid system with heat enhancement were produced. The idea starts with Choi [12] as he conducted an experimental study by added small amount of copper nanoparticle into conventional fluid with low thermal properties. The results shown that enhancement in term of thermal properties of the fluid system are achieved. In boundary layer problem, there is two well-known model as introduced by Buongiorno [13] and summarized by Tiwari and Das [14] where these two model are analyzed with different perspective. Buongiorno nanofluid model focused on the slip mechanisms that consist of inertia, Brownian diffusion, thermophoresis, diffusiophoresis, Magnus effect, fluid drainage, and gravity. Buongiorno [13] found that Brownian diffusion and thermophoresis are important factors to slip mechanisms in the nanofluids. Tiwari and Das [14] nanofluid model on the other hand focusing on the nanoparticle volume fraction and types of nanoparticle use in the fluid problem. This nanofluid model also discussed before by Khanafer et al. [15] in 2003 where this study considered the heat transfer performance of a nanofluid inside an enclosure. As for the nanofluid boundary layer problem with CHF boundary condition, Razi et al. [16] performed an experimental investigation of copper oxide nanoparticles in oil base fluid inside horizontal flattened tubes with different particle weight concentrations. Later, the laminar convective heat transfer in a two-dimensional micro-tube was studied numerically by Salman et al. [17] using four different types of nanoparticles and different nanoparticle size on CHF boundary condition. Recently, Sheikholeslamia and Zeeshan [18] studied the effect of Lorentz forces on water base nanofluid flow consist of copper nanoparticle in a permeable enclosure subjected to CHF boundary condition. The research field related to CHF boundary condition on nanofluid boundary layer flow was also conducted by consider different effects and flow properties.

Stagnation point flow defined as a flow with a point at the surface of boundary body that have zero local velocity and highest local pressure as derived from Bernoulli's equation. Due to interesting characteristic hold at the stagnation point region, lots of studies was conducted theoretically on the boundary layer flow [19–22]. Some study related to stagnation point flow with constant heat flux was conducted by Natalia and Pop [23] on steady mixed convection stagnation point flow over a vertical flat plate consist of both assisting and opposing flow regimes. Later, the studies continued by Mohammadiun et al. [24] on either a cylinder or flat plate of axisymmetric stagnation-point flow for incompressible fluid with constant heat flux. Different types of stagnation point flow then considered by Alizadeh et al. [25] where unaxisymmetric stagnation-point flow of an infinite stationary cylinder with non-uniform normal transpiration was conducted numerically at CHF boundary condition together with its heat transfer properties. There is also research conducted that consider nanofluid usage on their study at CHF boundary condition as performed by Li et al. [26] for laminar, incompressible, and steady oblique stagnation point flow of Cu—water nanofluid over a stretching/shrinking sheet with mass suction. There is some modification on heat flux model applied in this study where Cattaneo—Christov heat flux model was used in investigate the qualities of surface heat transfer. With the same heat flux model, Ahmad et al. [27] performed a study on micropolar fluid containing single-wall carbon nanotube and multi-wall carbon nanotube on a stagnation point flow with nonlinear stretching surface with slip effects. Some others interesting studies conducted on boundary layer flow was also consider an effect related to body force or any effects significantly influenced by gravitational force.

Effect that influence fluid flow does not only occurs under Earth gravitational where there is some effect that exist only under microgravity environment. G-Jitter is an effect that occur under microgravity environment where a fluctuating small gravitational acceleration is produced. Some studies related to g-jitter effect on boundary layer flow were conducted on different types of fluid with different fluid system orientation [28–32]. By considering CHF as the boundary condition, Shafie and Amin [33] performed a theoretical study on free convection flow induced by the g-jitter subjected to sphere CHF where it shown that convection flow was stronger at small Prandtl number. Recently, Amin et al. [34] continued the study with same boundary condition passes through an infinite length of vertical parallel plates of a double diffusion convection. The fluid system setup with the left wall of the plate is prescribed with constant heat flux while the right wall is maintained at a constant temperature. Later Hamdan et al. [35] performed a study on boundary layer flow near a three-dimensional stagnation point region induced by g-jitter effect where the mathematical model solved numerically using implicit finite different approach. There is a lot of boundary layer flow studies involving nanofluid that induced by g-jitter effect where various effects, boundary condition and model was considered [36–38]. There is also research conducted on stagnation point nanofluid flow induced by g-jitter but the problem is subjected to constant wall temperature [39]. From the literature, there is no theoretical study on nanofluid problem near a stagnation point region conducted subjected to heat flux boundary condition. Thus, motivated from the previous research and the important of CHF on stagnation point nanofluid flow, a

theoretical study was conducted mathematically on free convection nanofluid stagnation point flow induced by g-jitter effect with CHF boundary condition. The fluid problem was governed into a system of partial differential equation subjected to effect considered. The system of partial differential equation then solved numerically and discussed critically subjected to fluid behavior and thermal characteristic.

22.2 Problem Formulation

Under the microgravity environment, g-jitter effect was found to be significant on fluid system where a fluctuating gravitational field is produced. The effect causes a different behavior on fluid characteristic and heat transfer properties compared with normal gravitational field. Mathematically, the gravitational field with g-jitter effects depends on time t , is defined as,

$$g(t) = g_0[1 + \varepsilon \cos(\pi\omega t)] \quad (22.1)$$

where g_0 , ε and ω indicated the mean of the gravitational acceleration, amplitude of the gravity modulation and frequency of oscillation for the flow induced by the g-jitter. As for the fluid problem, a free convection flow near a three-dimensional stagnation point region is considered. A fluctuation gravitational field is occurring on the fluid system due to the consideration of g-jitter effect. Small amount of copper nanoparticles was added into the system with water were used as the base fluid. The constant temperature of the body is taken as T_w while T_∞ is the temperature of the fluid in the free stream. At the surface of the body, the stagnation point presented by nodal point \mathbf{N} , where the certain range values were specified by the geometrical shape of the body. By interpreting the fluid problem into three-dimensional orthogonal Cartesian system, the origin locates the nodal point where the x - and y - coordinates are measured along the body surface. The z - coordinate is than measured normal to the body surface and the physical model representation of stagnation point region in Cartesian coordinate system is presented in Fig. 22.1.

The fluid motion is described mathematically based on Newton's second law by Navier and Stokes, based on the movement of fluid element surface force and body force. Prandtl then come out with boundary layer theory which reduced the scope of the fluid flow to the surface of the boundary since the viscosity of the fluid is assumed to be significant only at the boundary. Based on nanofluid model summarised by Tiwari and Das [14] together with boundary layer and Boussinesq approximation, the fluid problem can be described into a system of partial differential equation such that,

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0, \quad (22.2)$$

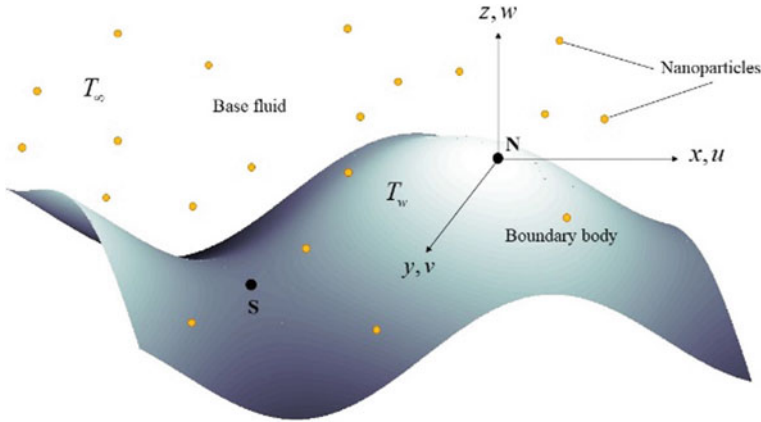


Fig. 22.1 Physical model representation of stagnation point region in Cartesian coordinate system

$$\rho_{nf} \left(\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} + w \frac{\partial u}{\partial z} \right) = \mu_{nf} \frac{\partial^2 u}{\partial z^2} + g(t)(\rho\beta)_{nf}ax(T - T_\infty), \quad (22.3)$$

$$\rho_{nf} \left(\frac{\partial v}{\partial t} + u \frac{\partial v}{\partial x} + v \frac{\partial v}{\partial y} + w \frac{\partial v}{\partial z} \right) = \mu_{nf} \frac{\partial^2 v}{\partial z^2} + g(t)(\rho\beta)_{nf}by(T - T_\infty), \quad (22.4)$$

$$\frac{\partial T}{\partial t} + u \frac{\partial T}{\partial x} + v \frac{\partial T}{\partial y} + w \frac{\partial T}{\partial z} = \alpha_{nf} \frac{\partial^2 T}{\partial z^2}, \quad (22.5)$$

subjected to constant heat flux boundary condition,

$$\begin{aligned} t < 0 : u = v = w = 0, T = T_\infty & \quad \text{for any } x, y \text{ and } z, \\ t \geq 0 : u = v = w = 0, k_{nf} \frac{\partial T}{\partial z} = -q_w & \quad \text{on } z = 0, x \geq 0, y \geq 0, \\ u = v = w = 0, T = T_\infty & \quad \text{as } z \rightarrow 0, x \geq 0, y \geq 0. \end{aligned} \quad (22.6)$$

Parameter u, v and w are the velocity component along the direction x, y, z axes respectively with T is the dimensional temperature parameter of the nanofluid. The thermophysical properties of the nanofluid such as density, dynamic viscosity, thermal expansion, thermal diffusion, and specific heat capacity at constant pressure are carried by with a term ρ, μ, β, α and C_p . The principle curvature denoted by parameter a and b were measured on the nodal point \mathbf{N} , at the $x = 0$ and $y = 0$ plane. The properties of the principal curvature follow $|a| \geq |b|$ with $a > 0$ and $c = b/a$. Here, c is defined as curvature ratio at \mathbf{N} taking a positive value with a range of $0 \leq c \leq 1$. Based on nanofluid model summarised by Tiwari and Das, the nanofluid correlation was defined from the Maxwell and Brinkman model for viscous fluid. The expressions of nanofluid correlation for viscous Newtonian nanofluid are,

Table 22.1 Thermophysical of the nanoparticles and base fluid

Physical properties	Water	Copper
Density ρ , (kgm^{-3})	997.1	8933
Specific heat capacity C_p , ($\text{Jkg}^{-1}\text{K}^{-1}$)	4179	385
Thermal conductivity k , ($\text{Wm}^{-1}\text{K}^{-1}$)	0.613	401
Thermal expansion $\beta \times 10^{-5}$, (K^{-1})	21	1.67

$$\begin{aligned} \mu_{nf} &= \frac{\mu_f}{(1 - \phi)^{2.5}}, \quad \alpha_{nf} = \frac{k_{nf}}{(\rho C_p)_{nf}}, \quad \frac{k_{nf}}{k_f} = \frac{(k_s + 2k_f) - 2\phi(k_f - k_s)}{(k_s + 2k_f) + \phi(k_f - k_s)}, \\ \rho_{nf} &= (1 - \phi)\rho_f + \phi\rho_s, \quad (\rho\beta)_{nf} = (1 - \phi)(\rho\beta)_f + \phi(\rho\beta)_s, \\ (\rho C_p)_{nf} &= (1 - \phi)(\rho C_p)_f + \phi(\rho C_p)_s. \end{aligned} \tag{22.7}$$

The nanoparticle volume fraction is defined by ϕ with k is defined as the thermal conductivity. The subscript f and s represent the fluid and solid component in the nanofluid mixture respectively. This study implies copper as the nanoparticle with water as the base fluid. Thus, the thermophysical properties of copper and water are shown in Table 22.1.

In a way to reduce the complexity of the problem, the system of partial differential equation will undergo semi-similar transformation technique that transform the equation into dimensionless system of equation. The semi-similar variables are introduced here as,

$$\begin{aligned} \eta &= Gr^{1/4}az, \quad \tau = \Omega t, \quad t = \nu a^2 Gr^{1/2}t^*, \\ u &= \nu a^2 x Gr^{1/2} f', \quad v = \nu a^2 y Gr^{1/2} h', \quad w = -\nu a Gr^{1/4}(f + h), \\ \Omega &= \frac{\omega}{\nu a^2 Gr^{1/2}}, \quad \theta = \frac{(T - T_\infty)}{(T_w - T_\infty)}, \quad Gr = \frac{g_0 \beta (T - T_\infty)}{a^3 \nu^2}, \end{aligned} \tag{22.8}$$

where ν is the kinematic viscosity of the fluid and Gr is the thermal Grashof number. The prime notation at the top of the function f and h indicate the differentiation with respect to η . Here, θ and Ω are the dimensionless variables for temperature, and frequency of oscillation. By using information in (22.7) and (22.8), the system of Eqs. (22.2)–(22.5) together with initial and boundary conditions (22.6) were transformed into dimensionless form such that,

$$C_1 f''' + C_2(f + h) f'' - C_2 f'^2 + C_3[1 + \varepsilon \cos(\pi \tau)]\theta = C_2 \Omega \frac{\partial f'}{\partial \tau}, \tag{22.9}$$

$$C_1 h''' + C_2(f + h) h'' - C_2 h'^2 + c C_3[1 + \varepsilon \cos(\pi \tau)]\theta = C_2 \Omega \frac{\partial h'}{\partial \tau}, \tag{22.10}$$

$$\frac{C_4}{C_5 Pr} \theta'' + (f + h)\theta' = \Omega \frac{\partial \theta}{\partial \tau}, \tag{22.11}$$

where Pr is Prandtl number and

$$C_1 = \frac{1}{(1 - \phi)^{2.5}}, \quad C_2 = 1 - \phi + \frac{\phi \rho_s}{\rho_f}, \quad C_3 = 1 - \phi + \frac{\phi(\rho\beta)_s}{(\rho\beta)_f},$$

$$C_4 = \frac{(k_s + 2k_f) - 2\phi(k_f - k_s)}{(k_s + 2k_f) + \phi(k_f - k_s)}, \quad C_5 = 1 - \phi + \frac{\phi(\rho C_p)_s}{(\rho C_p)_f}. \quad (22.12)$$

and dimensionless boundary condition,

$$f(\eta, 0) = f'(\eta, 0) = 0, \quad h(\eta, 0) = h'(\eta, 0) = 0, \quad \theta'(\eta, 0) = -1,$$

$$f' \rightarrow 0, \quad h' \rightarrow 0, \quad \theta \rightarrow 0, \quad \text{as } \eta \rightarrow \infty. \quad (22.13)$$

The dimensionless system of Eqs. (22.9)–(22.11) together with the boundary conditions (22.13) are then solved numerically using a finite difference approach. The solutions were analysed subjected to the dependent variables in term of skin friction coefficients and Nusselt number. The skin friction coefficient C_f and Nusselt number Nu are defined mathematically as,

$$C_{fx} = \mu_{nf} \left(\frac{\partial u}{\partial z} \right)_{z=0} / (\rho_f \nu^2 a^3 x),$$

$$C_{fy} = \mu_{nf} \left(\frac{\partial v}{\partial z} \right)_{z=0} / (\rho_f \nu^2 a^3 y),$$

$$Nu = -a^{-1} k_{nf} \left(\frac{\partial T}{\partial z} \right)_{z=0} / k_f (T_w - T_\infty). \quad (22.14)$$

The physical quantities in Eq. (22.14) were reduced into dimensionless form by taking the same semi-similar transformation procedure and becomes,

$$C_{fx}/Gr^{3/4} = f''(\tau, 0)/(1 - \phi)^{2.5},$$

$$C_{fy}/Gr^{3/4} = h''(\tau, 0)/(1 - \phi)^{2.5},$$

$$Nu/Gr^{1/4} = -(k_{nf}/k_f)\theta'(\tau, 0). \quad (22.15)$$

22.3 Results and Discussion

The dimensionless system of equation in (22.9)–(22.11) subjected to boundary condition (22.13) were solved numerically using implicit finite different scheme known as Keller box method. The solution is then presented graphically in term of profiles and physical quantities. Effects considered in the study such as amplitude of modulation

ε , frequency of oscillation Ω , curvature ratio c , together with nanoparticle volume fraction ϕ were analyzed and discussed critically. The effect of curvature ratio c , was analyzed and presented in Figs. 22.2 and 22.3 together with various values of amplitude of modulation ε . Figure 22.2 presents the skin friction coefficient on x - and y - direction with $c = 0$. It can be seen that for skin friction on y - direction, there is no changes with the increment of ε values. The value of $c = 0$ presenting cylindrical shape of geometrical body where the flow occurs here known as plane stagnation point flow cases. On the other hand, Fig. 22.3 shows the skin friction coefficient on both directions as $c = 1$ parameter value was induced. From the figure, the same amplitude values were noticed as the same ε was applied on the problem for skin friction on x - and y - directions. The stagnation point flow occurs here known as axisymmetric stagnation point flow case. The geometrical shape presented by $c = 1$ is spherical shape where it caused axisymmetric stagnation point flow case.

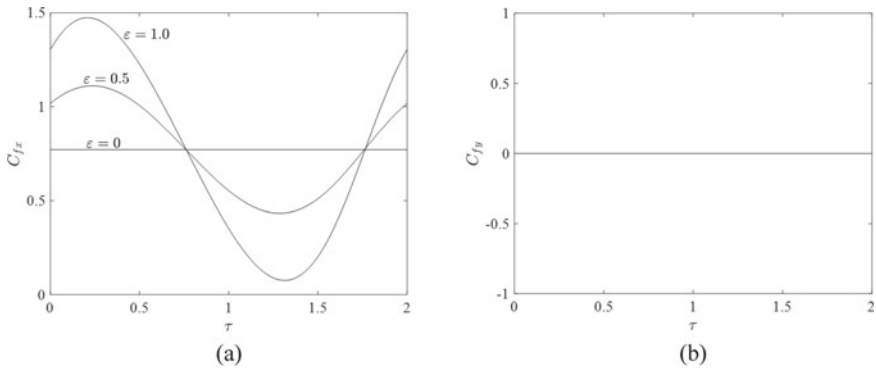


Fig. 22.2 The skin friction coefficient on **a** x - direction and **b** y - direction with various values of ε at $c = 0$

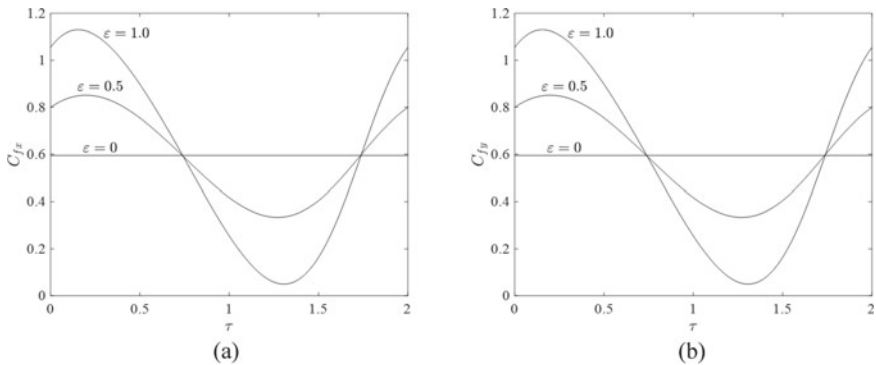


Fig. 22.3 The skin friction coefficient on **a** x - direction and **b** y - direction with various values of ε at $c = 1$

g-Jitter is an effect that occurs under microgravity environment that caused a fluctuation gravitational field. It is mathematically described subjected to two main effects which are amplitude of modulation ε and frequency of oscillation Ω . Figure 22.4 shows the velocity profile in x – direction with different sizes of Ω and various values of ε while Fig. 22.5 represents the temperature profile with the same cases by inducing different values of ε . From Figs. 22.4a and 22.5a, a fluctuation behavior is noticed for both profiles where g-jitter effect is found to be significant. In addition, the increase values of ε producing higher velocity profile compared to smaller values of ε while vice versa results were noticed for temperature profile. The bigger size of Ω that illustrated in term of profiles in Figs. 22.4b and 22.5b shown that increased of ε effect is found to be less significant.

The effect of nanoparticle volume fraction was illustrated in Fig. 22.6 in term of skin friction coefficient in x – direction on Fig. 22.6a and temperature profile in

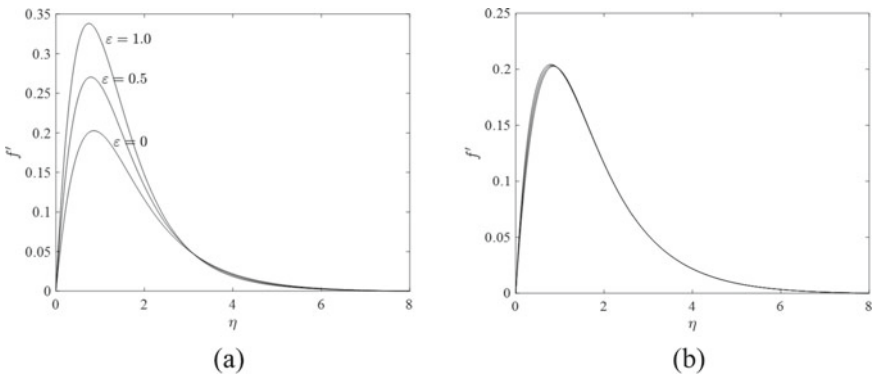


Fig. 22.4 The velocity profile on x – direction with **a** small size frequency of oscillation at $\Omega = 0.2$ and **b** bigger size frequency of oscillation at $\Omega = 5$ with various ε values

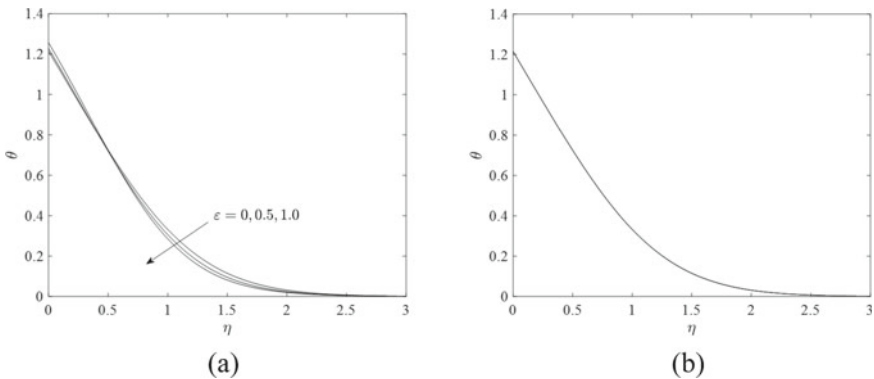


Fig. 22.5 The temperature profile with **a** small size frequency of oscillation at $\Omega = 0.2$ and **b** bigger size frequency of oscillation at $\Omega = 5$ with various ε values

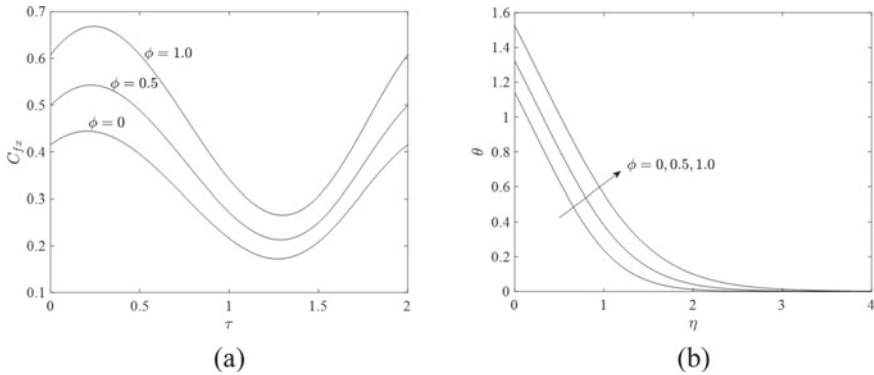


Fig. 22.6 The **a** skin friction coefficient on x – direction and **b** temperature profile with various values of ϕ

Fig. 22.6b. It is clearly can be seen that increases of ϕ increase both skin friction and temperature profile. The increases of skin friction are due to the additional of copper nanoparticle in the fluid system that producing extra resistance at the surface of the boundary body. Thus, additional friction will be produced at the body surface that will reduce the velocity of the fluid. On the other hand, enhancement of thermal properties was shown by Fig. 22.6b where the temperature profile increases with the increase of ϕ . Additional of copper nanoparticle that have higher thermal conductivity compared with conventional fluid shown an increment in term of temperature profile. Thus, a better thermal conductivity fluid system can be produced with consideration of nanofluid on the fluid system itself.

22.4 Conclusion

The unsteady free convection nanofluid flow near a three-dimensional stagnation point region induced by g-jitter effect with constant heat flux boundary condition has been studied theoretically. The fluid system was then govern in a form of partial differential equations and then solved numerically using implicit finite different approached. The results have been analyzed subjected to the effects considered in term of profiles and physical quantities. Based on the study, the effects considered significantly effecting the flow behavior and thermal characteristic that can be summarized as follow:

- The curvature ratio parameter determines the types of stagnation point flow,
- Increases of amplitude of modulation increase the velocity profile while decrease in the temperature profile,
- Bigger size of frequency of oscillation reduce the effect of amplitude of modulation on the fluid system,

- Increases of nanoparticle volume fraction increase the skin friction coefficient and temperature profile.

Acknowledgements The authors would like to acknowledge the Ministry of Higher Education Malaysia and Research Management Centre-UTM, Universiti Teknologi Malaysia (UTM) for financial support through vote numbers FRGS/1/2019/STG06/UTM/02/15, 08G33 and FRGS/1/2019/STG06/UTM/02/22

References

1. L. Prandtl, in *Proceedings of 3rd International. Mathematical Congress* (1904), pp. 484–491
2. G. Evans, R. Greif, *Int. J. Heat Mass Transf.* **34**, 2039 (1991)
3. N.L. Joseph, *J. Aerosp. Sci.* **25**, 194 (1958)
4. B.C. Sakiadis, *AIChE J.* **7**, 26 (1961)
5. S. Wu, *Ching. Appl. Sci. Res. Sect. A* **8**, 140 (1959)
6. M. Sudoh, K. Takuwa, H. Iizuka, K. Nagamatsuya, *J. Memb. Sci.* **131**, 1 (1997)
7. H.A.M. Al-Sharifi, L.A. Aziz, A.R.M. Kasim, M.Z. Salleh, and S. Shafie, in *National Conference on Postgraduate Research* (2016), pp. 886–891
8. J.H. Merkin, I. Pop, *Wärme- Und Stoffübertragung* **22**, 79 (1988)
9. B. Alazmi, K. Vafai, *Int. J. Heat Mass Transf.* **45**, 3071 (2002)
10. H. Ahmad, T. Javed, A. Ghaffari, *Thermophys. Aeromechanics* **24**, 115 (2017)
11. A.I. Alsabery, H. Saleh, I. Hashim, *Adv. Appl. Math. Mech.* **9**, 463 (2017)
12. S.U.S. Choi, in *Am. Soc. Mech. Eng. Fluids Eng. Div. FED* (1995)
13. J. Buongiorno, *J. Heat Transfer* **128**, 240 (2006)
14. R.K. Tiwari, M.K. Das, *Int. J. Heat Mass Transf.* **50**, 2002 (2007)
15. K. Khanafar, K. Vafai, M. Lightstone, *Int. J. Heat Mass Transf.* **46**, 3639 (2003)
16. P. Razi, M.A. Akhavan-Behabadi, M. Saeedinia, *Int. Commun. Heat Mass Transf.* **38**, 964 (2011)
17. B.H. Salman, H.A. Mohammed, A.S. Kherbeet, *Int. Commun. Heat Mass Transf.* **39**, 1195 (2012)
18. M. Sheikholeslami, A. Zeeshan, *Comput. Methods Appl. Mech. Eng.* **320**, 68 (2017)
19. T. Hayat, M. Hussain, M. Awais, S. Obaidat, *Int. J. Numer. Methods Heat Fluid Flow* **23**, 1155 (2013)
20. M. Zokri, S. Arifin, N. Syamilah, A.R. Mohd Kasim, M.Z. Salleh, *CFD Lett.* **12**, 1 (2020)
21. A.S. Abbassi, *J. Fluids Eng.* **131**, 1 (2009)
22. R. Subba, R. Gorla, V. Dakappagari, I. Pop, *Int. J. Heat Fluid Flow* (1993)
23. N.C. Ros, I. Pop, *Int. J. Heat Mass Transf.* **65**, 102 (2013)
24. H. Mohammadiun, A.B. Rahimi, A. Kianifar, *Sci. Iran.* **20**, 185 (2013)
25. R. Alizadeh, A.B. Rahimi, R. Arjmandzadeh, M. Najafi, *Alexandria Eng. J.* **55**, 1271 (2016)
26. X. Li, A.U. Khan, M.R. Khan, S. Nadeem, *Symmetry (Basel)*, **11**, 1 (2019)
27. S. Ahmad, S. Nadeem, N. Muhammad, M. Naveed, *J. Therm. Anal. Calorim.* **143**, 1187 (2021)
28. G.S. Kumar, *Glob. J. Eng. Technol. Adv.* **02**, 1 (2020)
29. I. Tlili, *Symmetry (Basel)*, **438**, 1 (2019)
30. S. Shafie, N. Amin, I. Pop, *Mech. Res. Commun.* **34**, 115 (2007)
31. M. Hidayad, A. Kamal, N.A. Rawi, M.R. Ilias, A. Ali, S. Shafie, *Univers. J. Mech. Eng.* **7**, 272 (2019)
32. M.J. Uddin, W.A. Khan, N.S. Amin, *PLoS One* **9**, e99384 (2014)
33. S. Shafie, N. Amin, *Malaysian J. Fundam. Appl. Sci.* **1**, 44 (2005)
34. W.Z.A. Amin, A.Q. Mohammad, S. Shafie, M. Qasim, *Sains Malaysiana* **49**, 1201 (2020)

35. F.R. Hamdan, M.H.A. Kamal, N.A. Rawi, A.Q. Mohamad, A. Ali, M.R. Ilias, S. Shafie, J. Adv. Res. Fluid Mech. Therm. Sci. **67**, 119 (2020)
36. B.S. Bhadauria, A. Singh, V. Kumar, Adv. Sci. Eng. Med. **10**, 707 (2018)
37. S. Agarwal, P. Rana, in Adv. Math. Sci. (2017), p. 020013
38. N.A. Rawi, A.R.M. Kasim, Z.M. Isa, A. Mangi, S. Shafie, Front. Heat Mass Transf. **8**, 1 (2017)
39. M.H. Ahmad Kamal, A. Ali, S. Shafie, Matematika **35**, 260 (2019)

Chapter 23

Spatial Decision Support System (SDSS) for Sustainable Business Location Selection of Purpose-Built Offices (PBO) in Malaysia



Jia Yun Heng, Edie Ezwan Mohd Safian, and Burhaida Burhan

Abstract The concept of sustainability has grown in popularity in the real estate industry around the world. On the global scale, the complex sustainability of real estate has been explored with the three main pillars of sustainability (environmental, social, and economic attributes) emerging. However, the characteristics of sustainable business location of purpose-built offices (PBO) have not been thoroughly investigated in Malaysia. In order to move toward sustainable business locations of PBO, Spatial Decision Support System (SDSS) is useful in assisting decision-makers in solving complex spatial decision problems. The SDSS is designed to provide a more systematic approach with geographic information for all manner of stakeholders. This paper aims to provide a SDSS for selecting sustainable business locations of PBO. In this paper, an Analytic Network Process (ANP) model and a Geographic Information System (GIS) are utilised as a SDSS to evaluate the sustainable business locations of PBO in Malaysia. Hence, combining ANP and GIS reveals and integrates both geographical data and decision makers' preferences to demonstrate a spatial data visualisation of sustainable business locations of PBO in the local context.

23.1 Introduction

In recent years, the sustainability issue has become a critical challenge across many disciplines. In Malaysia, the Green Building Index (GBI) was introduced in 2009 to promote sustainable building practices [1]. The sustainable building concept has grown in importance as a key element in achieving success in business throughout the years. Khamidi [2] stated that numerous office assessments have been established to evaluate the sustainability of different buildings since the 1990s. The location characteristics of a building have impacted its quality and sustainability of building through location evaluation [3].

J. Y. Heng (✉) · E. E. Mohd Safian · B. Burhan
Department of Real Estate Management, Faculty of Technology & Business Management,
Universiti Tun Hussein Onn Malaysia, 86400 Parit Raja, Johor, Malaysia

A sustainable location of the office for the business can maximise workplace quality of life while minimising the negative impacts in order to achieve the equilibrium of environmental, social and economic sustainability for businesses [4]. Pajones et al. [5] defined sustainable business location as a location for a business that sustains the resource land, supports long-term competitiveness, and meets the demands of the society. The sustainability of PBO's business locations should be studied in order to optimize the office space usage and improve the business performance. The complexity of spatial decision-making has increased, making it difficult for PBO stakeholders to make real estate decisions.

When it comes to distinct aims and objectives, spatial decision-making is often complex, multidisciplinary, and involves a variety of decision-makers [6]. To deal with decision problems related to spatial information, a computer-based support system known as Spatial Decision Support Decision (SDSS) combines the functionalities of DSS and the spatial analytic capabilities of GIS software [7]. SDSS is a system that uses both spatial and non-spatial datasets to address complex spatial decision problems. Real estate data must be collected, organized, and analyzed in order to solve complex spatial decision problems.

Because of the complexity of the spatial decision-making process, multi-criteria decision analysis (MCDA) can effectively solve such complicated problems. The integration of MCDA and GIS as a SDSS improves spatial decision-making by evaluating multiple attributes [6]. The ANP of MCDA has been used to solve location decision problems as it is a network that allows interdependence among clusters of elements and overcomes the limitations of linear hierarchic structures [8].

The ANP is an effective tool for quantifying the measurements for both spatial and non-spatial characteristics based on the judgments of decision makers. Therefore, the ANP method is more applicable to real world situations that rely on judgments and preferences [8]. The capabilities of GIS allow the combination of the preferences of decision-makers and geographical data to produce a standardized output in order to investigate the influence of spatial and non-spatial attributes in evaluating the location of a property.

The sustainable business location of PBO has increased the complexity in the evaluation and selection of PBO. The aim of this paper is to provide a SDSS to assist decision-makers in selecting a sustainable business location for a PBO in the local market. This paper employs the ANP and GIS as SDSS for sustainable business location selection of PBO. The ANP model is applied to investigate the preferences of stakeholders in the sustainable business location of PBO by taking into account both spatial attributes (transportation, accessibility and proximity) and non-spatial attributes (community identity and image). The results of the ANP analysis will be translated into GIS software, which will be used to analyze the sustainable business locations of PBO using GIS visualization techniques.

23.2 Spatial Decision Support System (SDSS)

As decision-making situations are often complicated, computer-based systems such as Geographic Information System (GIS), Decision Support System (DSS), and Spatial Decision Support System (SDSS) are required to solve complex decision problems. The DSS applications have grown in popularity across a wide range of disciplines for collecting, managing, and analysing massive amounts of data [6]. The DSS, on the other hand, does not make use of spatially referenced data. GIS is a computer technology that manipulates, stores, analyses, and displays geospatial information [9, 10]. GIS software is a key tool for assisting property stakeholders to understand the importance of spatial aspects of real estate [11].

In considering the spatial aspects of decision-making processes, it is necessary to develop a SDSS model that combines the DSS with GIS technology. Hence, SDSS has been designed and developed to support the spatial decision-making processes in situations where individuals are unable to address all the information efficiently and effectively without the aid of computer technologies [6]. SDSS is an integrated computer-based system that combines the functionalities of DSS and the spatial analytic capabilities of GIS software to assist decision-makers with complex spatial problems [7].

Over the past three decades, SDSS has grown and evolved following rapid technological developments [12]. The SDSS provides a framework for incorporating analytical and spatial modelling capabilities, spatial and non-spatial data management, domain knowledge, spatial display capabilities, and reporting capabilities into one system [6]. The benefits of SDSS allow stakeholders to analyze and evaluate a set of possible alternatives in order to make informed decisions.

23.3 Methodology

Reference [6] summarised that the SDSS has evolved from standalone desktop applications to web-based and service-based SDSS. Since SDSS evolved from DSS and GIS, the three major components of DSS (database, model base, and user interface) and GIS (database, user interface, and spatial data analysis and presentation capabilities) can be utilised in the development of SDSS. As a result of combining the components from both DSS and GIS, a database, a model base, and a user interface are three fundamental components of SDSS at the basic level. Figure 23.1 illustrates the architecture and decision-making process of SDSS.

Based on the spatial decision-making process shown in Fig. 23.1, the stakeholders define the problems as the input. The user interface component of SDSS facilitates the interaction between the stakeholders and the system in order to present visualisation and reports. Both spatial and non-spatial data can be organised, analysed, and displayed using the database management system's functionalities. As the core of the analytic model, the model management system is required to generate outputs using

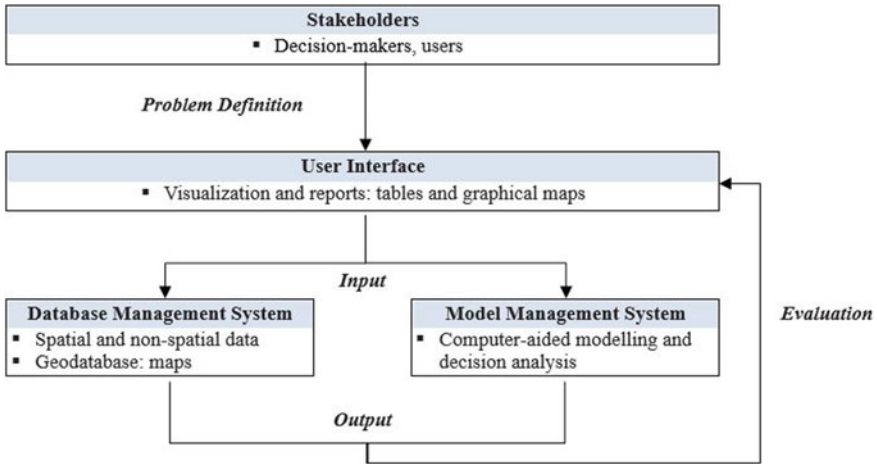


Fig. 23.1 The framework of spatial decision support system (SDSS)

computer-aided modelling and decision analysis. The outputs are presented in a way that enables the stakeholders to compare and evaluate the different alternatives.

During the data collection process, the researcher conducted a questionnaire-based survey to analyse the preferences of stakeholders towards sustainable business locations of PBO within the Golden Triangle Kuala Lumpur area. For this paper, the researcher used the two-stage cluster sampling method to determine the number of PBO that are represented as a whole. Two-stage cluster sampling is a method that combines cluster sampling and simple random sampling methods to identify each element that should represent the heterogeneity of the population [13].

For the first stage of sampling, the researcher chose four main streets which are Jalan Ampang, Jalan Sultan Ismail, Jalan Raja Chulan and Jalan Imbi in the research area using the cluster sampling method. By using a simple random sampling method, the researcher decided to randomly select 5 PBOs for every selected main street. However, only four PBO samples are available along Jalan Imbi. For this survey, 19 PBOs were chosen as samples. Accordingly, the researcher decided to apply the quota sample method to choose 10 occupants as the representatives of each PBO. The reason why quota sampling is the most appropriate method to be applied in this survey is that it allows the researcher to select the interest respondents for the research in investigating the stakeholders' preferences on the sustainable business locations of PBO. Thus, a total of 190 sets of questionnaires were distributed to the 19 PBOs respectively. The researcher has successfully collected the data needed with a high response rate of 82.63%.

The questionnaire survey was conducted to investigate the preferences of stakeholders towards sustainable business locations of PBO in the research area. The data was collected through a questionnaire that was given to the occupants of PBO. The researcher used a SDSS to evaluate the sustainable business locations of PBO in the

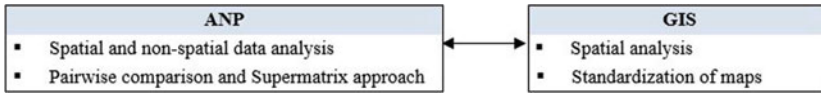


Fig. 23.2 The capabilities of ANP and GIS

research area. The weight of dependence for each sustainable business location characteristics of PBO was measured and compared using the ANP approach. The data collected from the questionnaire given according to the main streets was analysed and summarised by the researcher. The weight of dependence of each sustainable business location characteristic of PBO in the research area was determined through the analysis of the questionnaire survey.

The weight of dependence for each sustainable business location characteristic of PBO was calculated based on the pairwise comparison and supermatrix approach in the ANP application. The GIS software was used by the researcher to perform mapping analysis. The ANP data was translated into quantitative data that could be used as an input in the application of GIS. Hence, the visual illustrational results of the analysis can be used to observe the sustainable business locations of PBO. The SDSS was developed using the integration of ANP results with GIS technology to support the decision-making in selecting a desired sustainable business location of PBO. Figure 23.2 illustrates the capabilities of ANP and GIS in providing a SDSS for sustainable business locations of PBO.

23.4 Research Findings and Discussions

23.4.1 ANP Analysis

The ANP analysis was carried out to determine the interdependency of the sustainable business location characteristics of PBO. The ANP considers the interdependence of criteria in a network structure. Many decision problems cannot be organised in a hierarchical framework because they involve higher-level elements relying on lower-level elements [14]. Hsu [15] and Koc and Burhan [16] summarised that the ANP process consists of four general steps: identifying the problems, structuring the model, obtaining the priorities derived from pairwise comparisons and forming a supermatrix.

Pairwise comparisons were used to determine the weight of dependency for each sustainable business location characteristic of PBO. Using the supermatrix approach in the application of ANP, the results were used to generate the limiting values for each sustainable business location characteristic. The relative values for the sustainable business location characteristics of PBO were obtained from every column of the limit matrix are the same. Table 23.1 presents the limiting values of the sustainable business location characteristics of PBO according to the selected streets.

Table 23.1 Limiting values for sustainable business locations of PBO

Characteristics	Limiting values				Rank
	Jalan Ampang	Jalan Sultan Ismail	Jalan Raja Chulan	Jalan Imbi	
Location of commercial feature	0.027063	0.040359	0.039281	0.045315	9
Availability of transport options	0.067356	0.067597	0.054189	0.058278	2
Transportation distance	0.060621	0.057987	0.045447	0.054995	5
Vehicle flow	0.064666	0.043062	0.054423	0.037114	7
Efficiency of property market	0.054921	0.051691	0.071912	0.057221	3
Image/Branding of location	0.059190	0.058819	0.059819	0.039448	6
Access to amenities	0.058751	0.039559	0.050707	0.049539	8
Access to public transportation and terminal	0.071457	0.091208	0.059225	0.078856	1
Access to market	0.035973	0.049717	0.064997	0.079234	4

Table 23.1 shows the findings of the ANP analysis for the sustainable business location of PBO. The limiting values of sustainable business location characteristics are provided in the table above. The access to public transportation characteristic is the characteristic with the highest limiting value. The characteristic of availability of transport options has the second highest limiting value. It is followed by the efficiency of property market characteristic and the access to market characteristic. The limit matrix for all the streets reveals that the limiting values for the characteristics of transportation distance, vehicle flow, image/branding of location and access to amenities are on average level for all four streets. The location of commercial feature characteristic has the least limiting value.

In terms of the limiting values, it can be concluded that the characteristics of availability of transport options, access to public transportation and terminal, and efficiency of property market have significant limiting values for all the selected streets. According to the findings, there are transportation improvements that should be considered in the research area. Therefore, rather than developing new office spaces, both the government and private sectors should concentrate on improving existing office spaces. The findings of the ANP analysis showed that the limiting values generated could be incorporated into GIS software.

23.4.2 Spatial Analysis

Based on the findings from the ANP analysis, the researcher carried out a spatial analysis on the sustainable business location characteristics of PBO which consist of location of commercial feature, availability of transport options, transportation distance, vehicle flow, efficiency of property market, image/branding of location, access to amenities, access to public transportation and terminal, and access to market. The researcher implemented the GIS technology to support spatial analysis in order to produce useful information for sustainable business location decision-making in the research area. Each sustainable business location characteristic of PBO was represented as a map layer in the GIS. A map layer in a GIS database can be displayed and analysed independently or integrated with other map layers to produce a standardised output map layer. Therefore, the sustainable business location characteristics for all PBO samples can be visualised on a map. Figure 23.3 shows a visual of the ANP findings that have been transformed into a heatmap pattern through a GIS application.

Figure 23.3 shows visuals for the sustainable business location selection of purpose-built offices in the research area. The limiting value for each variable of sustainable business location that has been obtained is transformed into spatial data through GIS application. The pattern shows the heatmap pattern in which diameter areas are classified into 5 classes based on the limiting values in each variable for four areas. This pattern clearly shows that Jalan Raja Chulan, Jalan Sultan Ismail, and Jalan Imbi have large diameter patterns in which indicates that these three areas have potential locations and sustainable business opportunities in the context of purpose-built offices.

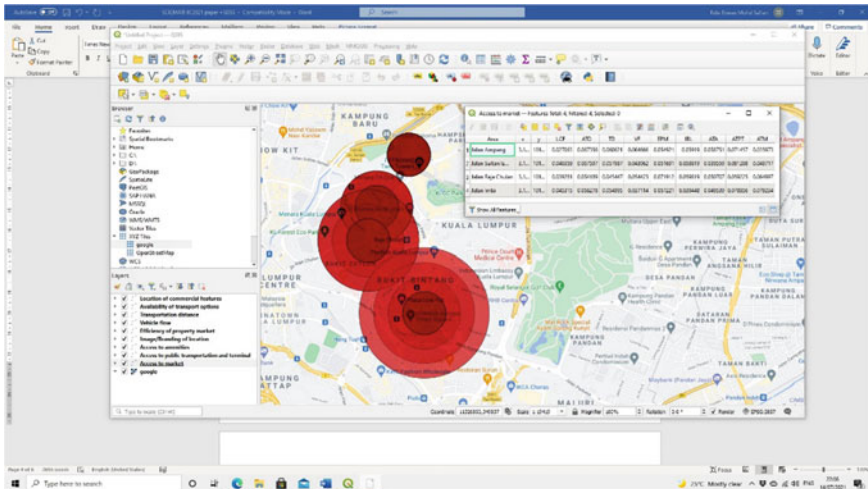


Fig. 23.3 Heatmap for sustainable business location selection of purpose-built offices

23.5 Conclusion

This paper utilised the ANP model and GIS software to develop a SDSS for selecting sustainable business locations of PBO in order to improve the performance of the PBO market in Malaysia. This paper bridges the gap by introducing a SDSS to the Malaysian real estate industry for selecting sustainable business locations of PBO. The results obtained in this paper will help to improve the sustainable development and performance of PBO buildings in Malaysia.

Acknowledgements The authors wished to acknowledge the Postgraduate Research Grant funded by the Office for Research, Innovation, Commercialization and Consultancy Management (ORICC) UTHM for supporting this study.

References

1. Y.M. Adnan, N.U. Aman, M.N. Razali, M.N. Daud, The implementation of green lease practices for office buildings in Kuala Lumpur, Malaysia. *Property Manage.* **35**(3), 306–325 (2017)
2. M.F. Khamidi, Development of building assessment tool for evaluation of purpose-built office life cycle management: benchmarking and assessment for environment performance, in *Conference on Sustainable Building South East Asia, Malaysia* (2007)
3. Guideline for Sustainable Building. Federal Ministry for the Environment, Nature Conservation, Building and Nuclear Safety (BMUB), Germany (2014)
4. A. Dalton, *Mapping the Sustainability of Small Business Locations* (Universtiy of the West of England, United Kingdom (UK), 2008)
5. M. Pajones, T. Steiner, N. Hackner-Jaklin, Development of a location-factor-matrix for sustainable business locations, in *Sustainable Development and Planning VIII, Austria* (2016)
6. R. Sugumaran, J. Degroote, *Spatial Decision Support System: Principles and Practices* (CRC Press, United States of America, 2011)
7. A. Eissa, Spatial decision support sytem for infrastructure resource allocation. Master thesis, University of Twente, Enschede, Netherlands (2013)
8. T.L. Saaty, *Theory and Applications of the Analytic Network Process: Decision Making with Benefits, Opportunities* (RWS Publications, Costs and Risks, 2009)
9. C.A. Rayed, Using GIS for modelling a spatial DSS for industrial pollution in Egypt. *Am. J. Geogr. Inf. Syst.* **1**(4), 100–104 (2012)
10. R.K. Sharma, D.P. Sharma, Review of Spatial Decision Support Systems in Resource Management. *Rev. Bus. Technol. Res.* **6**(1), 167–174 (2012)
11. R. Reed, C. Pettit, *Real Estate and GIS: The Application of Mapping Technologies*. Routledge of the Taylor and Francis Group, London and New York (2019)
12. P.B. Keenan, P. Jankowski, Spatial decision support systems: three decades on. *Decis. Support Syst.* **116**, 64–76 (2019)
13. S. Ahmed, *Methods in Sample Surveys* (The Johns Hopkins University, United States, 2009)
14. T.L. Saaty, L.G. Vargas, *Decision Making with the Analytic Network Process: Economic, Political, Social and Technological Applications with Benefits, Opportunities, Costs and Risks, Pittsburgh*. Springer Science + Business Media (2006)
15. P.-F. Hsu, Applying the ANP model for selecting the optimal location for an international business office center in China. *Asia Pac. Manag. Rev.* **15**(1), 27–41 (2010)
16. E. Koc, H.A. Burhan, Analytic network process in location selection and its application to a real life problem. *Int. J. Comput. Inf. Eng.* **10**(3), 507–512 (2016)

Part III
Statistics

Chapter 24

Investigation on the Effect of Household Income and Demographic Factors on Healthcare Cost in Perak



Khuneswari Gopal Pillay  and Wong Pui Yee

Abstract This paper focusing on the factor that effect the household income and demographic factors on healthcare cost in Perak as the increasing healthcare cost had brought many concerns to the community. Factors affecting the choice of hospital and healthcare is the main issue raise in this study. In the pilot study, 50 respondents were used to test for reliability using Cronbach's alpha. Result shows a reliability value of 0.785. Then, convenience sampling was used to collect response from 307 respondents who visited private and public hospitals in Perak. The data collected was used to determine the healthcare awareness of people in Perak by using the Chi-square test and cross-tabulation analysis. Meanwhile, LASSO and LASSO logistic regression were used to identify the factors affecting the choice of hospital. The better model in identifying factors affecting the choice of hospital was computed based on the accuracy, Brier score and AUC. Next, LASSO regression was applied to explore the effect of household income and demographic factors on healthcare cost in Perak. Based on the findings, people in Perak had high awareness of health insurance and personal healthcare. LASSO logistic regression was found to be the best model in identifying factors affecting the choice of hospital with a higher prediction accuracy, 0.7917 and lower Brier score, 0.3123. Variables affecting the choice of hospital and healthcare cost included education level, employment status, household income, ownership of household and having/not having an insurance policy. These findings could guide authority to improve the healthcare system for the benefits of people in Perak.

K. G. Pillay (✉)

Data Analytics, Sciences and Modeling Focus Group, Universiti Tun Hussein Onn Malaysia, 84600 Pagoh, Johor, Malaysia
e-mail: khuneswari@uthm.edu.my

K. G. Pillay · W. P. Yee

Department of Mathematics and Statistics, Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, 84600 Pagoh, Malaysia

24.1 Introduction

After decades of transformation, the healthcare industry has become one of the most important sectors in Malaysia to provide high-quality healthcare service [1]. According to the International Living 2019 Annual Global Retirement Index, Malaysia scored 95 out of 100 which ranked first place in providing quality healthcare service. [2]. Malaysia practices a two-tier healthcare system known as public and private. Public healthcare owned and run by the government to provide low-cost healthcare service to the people. Meanwhile, private healthcare owned and run by non-government organizations to provide faster and better service with the higher service cost [3]. The huge difference in fees charged by both type of hospitals causes a great burden to people who need to seek special treatment in private healthcare [4].

Starting from the year 2016, the healthcare cost keeps on increasing at an unforeseen speed [5]. This situation getting worse when non-communicable diseases (NCD) such as heart attacks and strokes have been increasing over the years along with the increase of life expectancy in Malaysia [6]. The risk of getting non-communicable diseases in the ageing group increases thus growing the demand for healthcare service [7]. Family with low financial stability may not have the ability to support the increasing demand.

A random survey was conducted in Perak as Perak has the highest ageing population and the healthcare expenditure for older population in Perak has increased significantly compared to previous years [8]. Pearson chi-square test was first conducted to determine the healthcare awareness of people. Logistic and LASSO logistic regression was carried out to identify the factors affecting the demand of hospital type. LASSO regression was then carried out to investigate the effect of household income and demographic factors on the healthcare cost in Perak. LASSO regression was chosen as it is well-suited for variable selection model to prevent overfitting [9].

24.2 Methodology

24.2.1 *Sampling Technique*

There were 10 private and 15 public hospitals in Perak. However, only 5 private and 3 public hospitals were selected to collect 300 respondents. Convenience sampling was used in sample selection in this study. Convenience sampling is non-probability sampling, where it is a method used to collect data from a group of people who are readily and easily approached [10].

24.2.2 Data Collection

There are five sections consist of the survey questionnaire used for collecting data from people who visited private and public hospitals in Perak. A total of 56 questions covered the demographic information, health status, health insurance, health awareness and a 5-point Likert scale regarding health insurance, personal health awareness, healthcare providers and Covid-19 awareness were generated. A face-to-face interview was carried out to collect response from both private and public hospital in Perak. Before that, 50 respondents were used to test for reliability during the pilot study.

24.2.3 Pilot Study

A pilot study is an important process of testing the feasibility and validity of research before the main study is conducted [11]. During the pilot study, 50 sets of survey questionnaires were distributed through social platform due to the Covid-19 pandemic. The reliability of the survey was tested using Cronbach's alpha which ranged from 0 to 1. A value nearer to 1 indicates greater reliability based on the rule of thumb.

24.2.4 Pearson's Chi-Square Test

Pearson's Chi-square test is a non-parametric test applied to discover the relationship between two categorical variables [12]. The null hypothesis is rejected when the p -value is less than 5% of significance level which indicates that there is a relationship between the two categorical variables. Hypothesis testing of Pearson's Chi-square test is stated as.

H_0 : The two categorical variables are independent.

H_1 : The two categorical variables are dependent.

24.2.5 Influential Observation

The observation was considered influential when the removal of this observation would cause a significant difference in the result. In this study, Cook's distance method was used to access the influential observations as in Eq. (24.1) [13].

$$D_i = \sum_{j=1}^n \frac{(\hat{Y}_j - \hat{Y}_{j(i)})^2}{(p+1)\hat{\sigma}^2} \quad (24.1)$$

where $i = 1, 2, \dots, n$ and \hat{y}_j is the j th fitted value and $\hat{y}_{j(i)}$ is the j th fitted value without i th observation, p is the number of coefficients in the regression model and $\hat{\sigma}$ is the estimated variance from the fitted value. Any observations that exceed the cut-off value ($4/n$) will be removed.

24.2.6 Logistic Regression

Logistic regression is a statistical technique to determine the relationship between dichotomous dependent variables and independent variables [14]. In this study, the dependent variable was the choice of hospital with 0 represents a private hospital and 1 represents a public hospital. In the beginning, the dataset was split into 90% of training data and 10% of test data. The independent variable involved in this study included qualitative and quantitative data. The dependent variable is expressed in terms of odds with the changes in the independent variable. By adding natural logarithm to the odds, the log odds (logit) of the dependent variable can be expressed as in Eq. (24.2) [15].

$$\text{Logit}(p) = \log \log \left(\frac{p}{1-p} \right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k \quad (24.2)$$

Backward elimination was used in selecting the significant variables to obtain a model with high accuracy based on Akaike Selection Criterion (AIC) [16]. It has a better predictive power with the wider range of possible predictor combinations. It begins with all the variables included in the full model and ends when the lowest AIC value obtain.

24.2.7 Model Diagnostic and Evaluation

The best-predicted model computed from logistic and LASSO logistic regression were evaluated based on the accuracy, Brier score and Area Under the Curve (AUC) [17]. The accuracy of the model was obtained through the classification table of the observed and predicted outcome. The Brier score is a cost function used to access the accuracy of probability predictions of outcome with the formula [18], in Eq. (24.3)

$$BS = \frac{1}{N} \sum_{t=1}^N (f_t - o_t)^2 \quad (24.3)$$

where N is the number of items, f_t is the correct forecast probability and o_t is the event outcome which is equal to 1 or 0. Next, a Receiver Operating Curve (ROC) is a classification accuracy graph that illustrates the sensitivity and specificity of every possible cut-off [19]. The area under the curve of ROC measures the performance of the forecast and ranged from 0 to 1. A model with a low Brier score, high accuracy and AUC was chosen as the better in predicting the choice of hospital.

24.2.8 Data Transformation

Data transformation could be used to transform the nature of variable when the distribution of data is greatly skewed [20]. Logarithm transformation was used in this study to transform the non-normal data with a p -value less than 0.05 to reduce the error and variability. Logarithm transformation is used due to its ability to transform highly skewed variable into a more normalized dataset especially dataset that contains outlying observations [21].

24.2.9 Least Absolute Shrinkage and Selection Operator (LASSO) Regression

The Least Absolute Shrinkage and Selection Operator (LASSO) regression is a penalized regression method that applied regularization to eliminate variables with unstable estimates to zero [22]. In the beginning, the dataset was split into training (80%) and test (20%) data in cross-validation to obtain the minimum lambda, λ . Next, training data was used to build the best LASSO model with the minimum lambda. The coefficients of each significant variable were then obtained and lastly, the Mean Square Error of Prediction, MSE(P) of the best model was computed using the test data. The formula of MSE(P) as shown in Eq. (24.4) [23].

$$MSE(P) = \frac{1}{t} \sum_{i=1}^t (\hat{Y}_t - Y_t)^2 \quad (24.4)$$

where $i = 1, 2, \dots, n$ is the sample size, \hat{Y}_t is the estimated y of test value and Y_t is the actual test value.

24.2.9.1 Goodness-of-Fit Test

The goodness-of-fit test is a diagnostic test that commonly used in determining how well a model fits the specific distribution [24]. The homoscedasticity and linearity of the fitted model were accessed through the residual plot and Q-Q plot.

The normality tests such as the Shapiro–Wilk test, Kolmogorov–Smirnov test and Anderson–Darling test were also applied in this study. The null hypothesis is accepted when the p -value is greater than 0.05, this indicates that the fitted model follows a normal distribution. The hypothesis testing for the three normality tests is as follow:

H_0 : The fitted model is normally distributed.

H_1 : The fitted model is not normally distributed.

24.3 Results and Discussion

24.3.1 Reliability Test

A reliability test was carried out using Cronbach's alpha with 36 Likert-scale typed questions in the questionnaire in Microsoft Excel. The result shows a value of 0.785 of internal consistency which indicates that the internal consistency of the items is in an acceptable range based on the rule of thumb. The scale used in the survey questionnaire is reliable.

24.3.2 Descriptive Analysis

The variables used in this study included gender, age, marital status, education level, employment status, household income, geographical area, ownership of household, having/not having a disease, frequency of health problem, type of hospital, number of dependents and having/not having an insurance policy. There were a total of 307 respondents participated in this survey. Among the 307 respondents, 58% of the respondents were female while 42% were male. The age group of 20–30 years old had the highest percentage of respondents which was 26.4% while the age group of more than 60 years old (>60) recorded the lowest percentage of participation which only consist of 15%. Meanwhile, 44.6% of the respondents were having high school education level and most of the respondents were married with a percentage of 70.7%.

24.3.3 Pearson's Chi-Square Test

Based on the result from Table 24.1, the p -value for all the variables in health insurance and personal health is less than a 5% significance level. These indicates that there is a significant relationship between health insurance and personal healthcare

Table 24.1 Pearson’s Chi-square test

	Variables	<i>p</i> -value	Decision
Health Insurance	Health insurance is important	< 0.001	Reject H_0
	Health insurance is affordable to you	< 0.001	Reject H_0
	I will still purchase health insurance regardless of how much income I earned	< 0.05	Reject H_0
	The coverage of insurance is enough to cover all the medical expenses	< 0.01	Reject H_0
	Satisfied with your current insurance policy	< 0.001	Reject H_0
	I have enough health insurance knowledge before I purchase the policies	< 0.001	Reject H_0
Personal Health	It is important to do a regular medical check-up	< 0.1	Reject H_0
	Aware of various symptoms of diseases	< 0.05	Reject H_0
	Self-prescribing when getting illness	< 0.001	Reject H_0

with spending in healthcare. Health insurance shows the highest awareness as most of the variables has lower *p*-value and correlated to the spending in healthcare.

24.3.4 Data Transformation

Before producing the best predicted LASSO model, the outliers and influential points were removed. Based on the assumptions of LASSO regression, outliers needed to be removed before model building while in logistic regression the removal of outliers is not required. There was a total of 176 observations left for LASSO model building. Data transformation was then carried out for the continuous variables, household income and spending in healthcare as both variables were not normally distributed (*p*-value < 0.05) after tested with normality test. After the log transformation, variable household income resulted in a smaller *p*-value while variable spending in healthcare resulted in a larger *p*-value as shown in Table 24.2. Both the transformed income and spending data were used in the model building process.

Table 24.2 Shapiro–Wilk test before and after transformation

Variable	<i>P</i> -value	
	Before transformation	After transformation
Household income	0.00002	7.975×10^{-10}
Spending	3.405×10^{-14}	8.451×10^{-5}

24.3.5 Logistic Regression

At the beginning of the analysis, Cook’s distance was used to detect and remove the influential points. There was a total of 248 observations left for model building after the removal of influential points. The qualitative data were coded, and the dataset was split into training (90%) and test data (10%). The model of logistic regression was built using the training set and backward elimination was used to remove the insignificant variables. The best-predicted model is as shown:

$$\begin{aligned}
 P_r(Y = \text{Public hospital}) = & 18.71 - 17.10 * \text{Education}_{\text{HighSchool}} - 16.16 * \text{Education}_{\text{STPM}} \\
 & - 17.75 * \text{Education}_{\text{Diploma/College}} - 17.08 * \text{Education}_{\text{Bachelor's Degree}} \\
 & - 35.53 * \text{Education}_{\text{Master's Degree or Higher}} + 5.9460 * \text{Education}_{\text{Others}} \\
 & - 0.6551 * \text{Income} + 1.3270 * \text{Area}_{\text{Rural}} - 0.9770 * \text{House}_{\text{Rent}} \\
 & - 2.0290 * \text{House}_{\text{Others}} - 0.0007 * \text{Spending} \\
 & - 0.6551 * \text{Insurance}_{\text{Yes}} + 0.3218 * \text{Dependents}
 \end{aligned}$$

24.3.6 LASSO Logistic Regression

Next, the training set is used in building a LASSO logistic model. Figure 24.1 shows the optimal lambda obtained from the Mean Square Error plot which was 0.0189.

The best predicted LASSO logistic model is

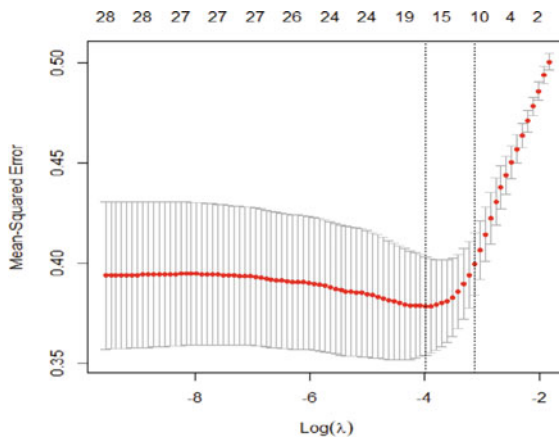


Fig. 24.1 Cross-validated MSE plot of the best-predicted model

Table 24.3 Comparison of predictive performance between two best-predicted model

Model	Accuracy	Brier score	AUC
Logistic model	0.75	0.50	0.7917
LASSO logistic model	0.7917	0.3123	0.7570

$$\begin{aligned}
 P_r(Y = \text{Public hospital}) = & 0.9259 - 0.0499 * \text{Age}_{41 \text{ to } 50} - 0.4285 * \text{Education}_{\text{Diploma/College}} \\
 & - 0.0983 * \text{Education}_{\text{Bachelor's Degree}} - 2.1253 * \text{Education}_{\text{Master's Degree or Higher}} \\
 & + 2.9859 * \text{Education}_{\text{Others}} + 0.0418 * \text{Marital}_{\text{Married}} \\
 & - 0.3447 * \text{Marital}_{\text{Divorced}} - 0.2510 * \text{Employment}_{\text{Self-employed}} \\
 & - 0.0002 * \text{Income} + 0.7822 * \text{Area}_{\text{Rural}} \\
 & - 0.0445 * \text{House}_{\text{Rent}} - 1.0546 * \text{House}_{\text{Others}} \\
 & - 0.0003 * \text{Spending} + 0.2307 * \text{Dependent} \\
 & + 0.1546 * \text{Health Problem}_{\text{Long term}} - 0.3331 * \text{Insurance}_{\text{Ye}}
 \end{aligned}$$

24.3.7 Comparison Between Logistic and LASSO Logistic Model

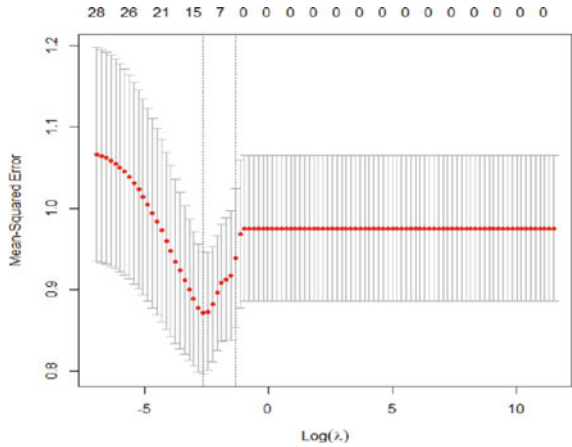
Table 24.3 shows the comparison of the predictive performance of both logistic and LASSO logistic model based on accuracy, Brier score and area under the curve (AUC). Based on the result, LASSO logistic model was chosen as the better model to predict the choice of hospital with higher predictive accuracy and lower Brier score.

Based on the result of LASSO logistic model, the significant variables which affect the choice of hospital included age, marital status, education level, employment status, household income, geographical area, ownership of household, frequency of health problem, spending in healthcare, number of dependents and having/not having an insurance policy. Community who bought insurance and with higher education level, higher income, higher spending in healthcare tend to choose private hospital. Meanwhile, community who lived in rural area, with long-term health problem and with higher number of dependents tend to choose public hospital.

24.3.8 LASSO Regression

In the formation of the LASSO model, 176 observations were included, and the dataset was split into training (80%) and test set (20%). The LASSO model was built using the training set and the optimal lambda was obtained through the Mean Square Error plot which was 0.07222 as shown in Fig. 24.2. The Predicted Mean-Square Error (MSE(P)) showed a value of 0.7370 for the best predicted LASSO model which

Fig. 24.2 Cross-validated MSE plot of best predicted LASSO model



is the lowest MSE(P) obtained to determine the effect of demographic factors and household income on healthcare cost in Perak.

The best predicted LASSO model is

$$\begin{aligned}
 \text{Spending} = & 5.2943 + 0.2126 * \text{Gender}_{\text{Female}} - 0.0032 * \text{Education}_{\text{Master's Degree or Higher}} \\
 & + 0.2330 * \text{Employment}_{\text{Self-employed}} - 0.0713 * \text{Employment}_{\text{Others}} \\
 & + 0.0549 * \text{Income} - 0.3611 * \text{Home}_{\text{Others}} + 0.1755 * \text{Diseases}_{\text{Yes}} \\
 & - 0.2144 * \text{Hospital}_{\text{Public}} + 0.4386 * \text{Insurance}_{\text{Yes}}
 \end{aligned}$$

Based on the result, it was found that gender, education level, employment status, ownership of household, choice of hospital, household income, having/not having any disease and having/not having an insurance policy have a significant effect on healthcare spending. Female was found to have a higher spending in healthcare. Community who was self-employed, with higher income, had a disease and bought an insurance were found to have a higher healthcare cost.

24.3.9 Goodness-Of-Fit Test

The goodness-of-fit test was carried out to examine how well the best predicted LASSO model fits the distribution. Residual plot and Q-Q plot were used to test for homoscedasticity and linearity of the fitted model as shown in Fig. 24.3.

Based on Fig. 24.3, the residuals had a constant variance, and the data points were normally distributed along the axis. Normality test such as Shapiro–Wilk, Kolmogorov–Smirnov and Anderson–Darling test were conducted as shown in Table 24.4. The *p*-value of three normality tests proved that the best-predicted model was

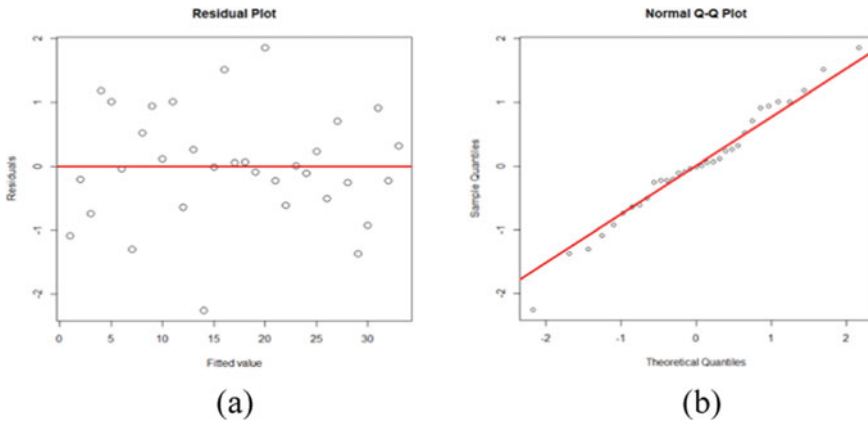


Fig. 24.3 Residual plot (a) and Q-Q plot (b) for the fitted model

Table 24.4 Normality test for best predicted LASSO model

Normality test	<i>P</i> -value
Shapiro–Wilk	<0.05
Kolmogorov–Smirnov	<0.05
Anderson–Darling	<0.05

normally distributed since all of them were greater than 0.05. Therefore, it can be concluded that the test data fitted the best LASSO model well.

24.4 Conclusion

People in Perak have a high awareness of health insurance and personal healthcare in terms of spending when most of the variables in health insurance and personal healthcare were correlated to the spending in healthcare. Generally, education level, employment status, ownership of household, household income and having/not having an insurance policy were found to affect the choice of hospital and spending in healthcare. Hence, it can be said that financial-related factors are significantly affecting the healthcare cost in Perak.

Followed by the high awareness of health insurance, the insurance company should implement more policies to fulfil the needs of people in Perak. Also, the government and policymakers should design more policies and allocate more budgets to improve the healthcare system for the benefits of people from different background especially with the current Covid-19 pandemic which gives a higher burden to people in prevention spending [25].

Some recommendations for future research are to include a larger sample size for better result. Next, future researchers are recommended to include more factors in the

research such as socio-economic factors to have a more comprehensive understanding of their effect towards the choice of hospital and healthcare cost. Finally, this study only focused on the effect of household income and demographic factors on the healthcare cost of people in Perak. Hence, the findings and result might only be useful for authorities in Perak on the decision-making. Future research on other states in Malaysia can be conducted for the overall picture of effect in Malaysia's healthcare service.

Acknowledgements The authors would like to thank the Universiti Tun Hussein Onn Malaysia (UTHM) for giving the chance to carry out this research and funding. The authors are also extremely thankful to the reviewers for their motivating remarks.

References

1. C.P. Yu, D.K. Whynes, T.H. Sach, Equity in health care financing: the case of Malaysia. *Int. J. Equity Health* **7**(1), 1–14 (2008)
2. M.D. Baltimore, The 6 Countries with the Best Healthcare in the World 2019. Retrieved on 10 June 2019, from https://www.prweb.com/releases/the_6_countries_with_the_best_healthcare_in_the_world_2019_internationalliving_com/prweb16049883.htm
3. N. Rahman, M. Husin, K. Dahian, K.M. Noh, R. Atun, S. Sivasampu, Job satisfaction of public and private primary care physicians in Malaysia: analysis of findings from QUALICO-PC. *Hum. Resour. Health* **17**(1), 1–10 (2019)
4. S. Muniapan, Healthcare costs in Malaysia. Retrieved on 25 May 2019, from <https://www.centre.my/post/healthcare-costs-in-malaysia>
5. T. Taner, J. Antony, Comparing public and private hospital care service quality in Turkey. *Int. J. Health Care Quality Assurance Incorpor. Leadership Health Services* **19**(2–3), 127–133
6. G.C. Brown, Living too long: the current focus of medical research on increasing the quantity, rather than the quality, of life is damaging our health and harming the economy. *EMBO Rep.* **16**(2), 137–141 (2015)
7. S. Maslina, A.K. Aizatu, Health care expenditure of ageing population in Perak, Department of Statistics Malaysia (2020)
8. J. Ranstam, J.A. Cook, LASSO regression. *J. Br. Surg.* **105**(10), 1348–1348 (2018)
9. N. Showkat, H. Parveen, Non-probability and probability sampling. *Natl. J. Multidisciplinary Res. Dev.* **2**(1), 55–60 (2017)
10. M. Mansor, N.Z. Harun, Health issues and awareness, and the significant of green space for health promotion in Malaysia. *Proc. Soc. Behav. Sci.* **153**, 209–220 (2014)
11. Z.A. Hassan, P. Schattner, D. Mazza, Doing a pilot study: why is it essential? *Malaysian Family Physician: Official J. Acad. Family Physicians Malaysia* **1**(2–3), 70–73 (2006)
12. M.L. Mchugh, The Chi-square test of independence lessons in biostatistics. *Biochemia Medica* **23**(2), 143–149 (2013)
13. R.D. Cook, Detection of influential observation in linear regression. *Technometrics* **19**(1), 15–18 (1977)
14. H.A. Park, An introduction to logistic regression: from basic concepts to interpretation with particular attention to nursing domain. *J. Korean Acad. Nurs.* **43**(2), 154–164 (2013)
15. S. Sperandei, Understanding logistic regression analysis. *Biochemia Medica* **24**(1), 12–18 (2014)
16. Q. Li, Logistic and SVM credit score models based on lasso variable selection. *J. Appl. Math. Phys.* **7**(05), 1131 (2019)

17. Q.G. Wang, J.J. Koval, C.A. Mills, K.I. Lee, Determination of the selection statistics and best significance level in backward stepwise logistic regression. *Commun. Stat. Simul. Comput.* **37**(1), 62–72 (2007)
18. M.S. Roulston, Performance targets and the brier score. *Meteorol. Appl. J. Forecast. Pract. Appl. Training Tech. Model.* **14**(2), 185–194 (2007)
19. R.A. Giancrifofaro, L. Salmaso, Model performance analysis and model validation in logistic regression. *Statistica (Bologna)* **63**(2), 375–396 (2007)
20. F. Changyong, W. Hongyue, L.U. Naiji, C. Tian, H.E. Hua, L.U. Ying, Log-transformation and its implications for data analysis. *Shanghai Arch. Psychiatry* **26**(2), 105 (2014)
21. J. Osborne, Notes on the use of data transformations. *Pract. Assess. Res. Eval.* **8**(1), 6 (2002)
22. Y. Mansiaux, F. Carrat, Detection of independent associations in a large epidemiologic dataset: a comparison of random forests, boosted regression trees, conventional and penalized logistic regression for identifying independent factors associated with H1N1pdm influenza infections. *BMC Med. Res. Methodol.* **14**(1), 99 (2014)
23. M.M. Amiri, L. Tapak, J. Faradmal, J. Hosseini, G. Roshanaei, Prediction of serum creatinine in hemodialysis patients using a kernel approach for longitudinal data. *Healthcare Inf. Res.* **26**(2), 112–118 (2020)
24. A. Maydeu-Olivares, C. Garcia-Forero, Goodness-of-fit testing. *Int. Encycl. Edu.* **7**(1), 190–196 (2010)
25. A. Knopf, Telepsychiatry coming into its own with COVID-19. *Brown Univ. Child Adolesc. Psychopharmacol. Update* **22**(5), 1–3 (2020)

Chapter 25

Classification Tree of Breast Cancer Data with Mode Value for Missing Data Replacement



Nur Atiqah Hamzah, Sabariah Saharan, and Khuneswari Gopal Pillay

Abstract Classification is one of the methods in data mining that can be used to group the data into class attributes based on similarities shared in the data. The decision tree has been widely used in many field industries as its simplicity of computation is suitable to be used to analyse various kinds of data available. However, missing readings in the data set often being omitted as classification cannot be done with missing readings in the data. Therefore, this study aims to carry out classification without deleting the missing part of the data. In this research, the performance of the decision tree will be used in analysing the Breast Cancer data. The breast cancer data were classified into two groups of recurrence and non-recurrence cancer and will be discussed. The patients' age, breast, breast quad, menopause, tumour size, involve nodes, node caps, degree of malignant, and irradiation are the attributes used for the classification process. The decision rules for the decision tree is obtained and provide more understandings based on the tree formed. The accuracy of the decision tree for both sets of data were presented in this study.

25.1 Introduction

25.1.1 Background of Study

Cancer is one of the leading causes of human death in the world and breast cancer has one of the highest rankings among other cancers. In Malaysia, about 17.3% of cancer patients suffered from breast cancer as stated by the World Health Organization in 2018. Breast cancer is the top cancer in women both in developed and developing countries. In this study, breast cancer data will be used to identify how cancer can be grouped into recurrence and non-recurrence cancer. Besides, the missing data will

N. A. Hamzah (✉) · S. Saharan · K. G. Pillay
Department of Mathematics and Statistics, Faculty of Applied Science and Technology, Universiti Tun Hussein Onn Malaysia, Batu Pahat, Johor, Malaysia

also be furthered analyze as there must be differences resulting in analysis that might contribute to significant results. These processes involved a few methods.

Data mining is the process of gaining useful information from large data repositories to find unexpected relationships and to summarise the data in novel ways to make the data understandable and useful to the data owner [1–3]. In data mining, the databases that have always been used comprise millions of records and thousands of variables. The existence of too many predictor variables usually complicates the process to model a relationship with a response variable. Thus, data mining techniques are an approach to find understandable solutions [4]. The well-known data mining techniques are anomaly detection, association rule learning, classification, clustering, and regression [5]. The common data mining tasks are description and summarization, concept descriptions, segmentation, classification and case-based reasoning, prediction, and dependency analysis [2, 6].

The purpose of deploying data mining techniques is to discover crucial patterns from datasets and also to provide the capabilities of predicting the outcome of a future observation. However, the relationships that exist in the data tend to be unclear as the amounts of information are too big or the types of relationships are very difficult to imagine [7]. Data mining is an extension of statistical methods that are known to be a technique of increasing the productivity of people trying to build a predictive model [8].

Classification is a technique that maps data into predefined groups or classes. By using the classification method, the class for the specific data used can be determined. Classification is one of the well-known methods in data mining that has been widely used in many sectors, which is known as the method with a simple computation technique. The data repositories that provide real-world data applications can always be used for classification purposes. Classification is a popular task in knowledge discovery and upcoming plans. It has been known as intelligent decision making as it examines sample data and also predicts the future behaviour of the data [9]. Classification is known as a technique of data analysis that results in discovering data classes [10]. The classification process can be applied to various types of data to predict categorical class labels [11].

Every decision tree consists of the root nodes, internal nodes, and leaf nodes whereby each of the nodes represents the classification rules of the tree. The classification tree method is also used in the medical field. According to [12], the classification tree helps in the analysis of cardiogram data for fetal distress determination. [13] used data mining approaches including a classification tree to detect heart disease while [14] used classification on breast cancer data. To accurately forecast the occurrence of heart disease with a condensed number of attributes, the study of J48 Decision Tree, Naive Bayes, and Bagging algorithm were applied. The cross-validation method was employed to measure the performance of the model. Decision trees can handle high dimensional data and have convenient accuracy. The classification always starts from the root node by testing the attribute specified at the root node and the value of the attribute is allocated at the tree branch [15].

Classification and regression tree or knowns as CART is one of the decision trees that can be used to classify the data [16]. It characterized by the fact it constructs

binary trees, in which each internal nodes have two outgoing edges. Its ability to generate regression trees where the leaves predict a real number instead of a class [17]. CART can be used to classify both numerical and categorical variables. Besides, CART identify the most significant variables for each decision tree, thus, the non-significant ones will not be displayed in the tree. Besides, CART also can handle outliers. Therefore, in this study, the decision tree built was based on the CART algorithm using Rstudio.

25.1.2 Objective of Study

The main objective of this research is to identify the performance of the decision tree to classify Breast Cancer into two specified classes of Non-recurrence and Recurrence cancer. Cancer is a disease of organs and tissues of the body that can affect any part of the body [13, 14]. Breast cancer is the second most leading cancer that happened in women compared to other cancers. Besides, this cancer is fatal in under half of all cases. Breast cancer occurs when cells within the breast duct and lobules become cancerous. In this study, data mining will be applied to classify breast cancer data. Missing data always occurred in an analysis [18]. As the data contains missing readings, the mode of each variable will be used to replace the missing ones. Therefore, there will be two decision trees represented the breast cancer classification process. The performance of the decision tree is further discussed in this paper. The classification of this data can be useful to predict the characteristics of each data to be in the same class based on the decision rules produced.

25.2 Methodology

In this study, the data used were Breast Cancer data provided by UCI Machine Learning. This breast cancer domain was obtained from the University Medical Centre, Institute of Oncology, Ljubljana, Yugoslavia. This data set includes 201 instances of one class and 85 instances of another class, which are known as non-recurrence and recurrence classes respectively. The data consists of 9 attributes which are nominal and categorical attributes. The attributes are age, breast, breast quad, menopause, tumor size, involved nodes, node caps, degree of malignant and irradiation. The data shows related attributes to the patients who previously had cancer and the presence of cancer again after treatments. The methods used for this study are discussed in depth in next sections.

25.2.1 Data Cleaning

Data cleaning is the process that takes place after data collection, which includes outlier detection, rule-based data cleaning, data transformation, and data deduplication [19]. Many researchers in various areas conducted data analysis without comprehending the importance of data cleaning as the analysis results can be totally amiss [20].

The breast cancer data had missing values despite the attributes which in categorical data using string characters. [3] stated that missing data can cause problems in the training and classification phase. The data training is a process to train the algorithm or machine learning model in order to predict the outcome of the proposed method. Meanwhile, the testing data is a number of data used to test and validate the performance of the methods used. Despite using the data directly, this study used train and test data which consisted of 90% and 10% of total number of data respectively. Meanwhile, there are a few causes for non-availability that exist in a set of data which are:

1. Equipment malfunction
2. Deletion due to inconsistency with other recorded data
3. Non-entry of data due to misunderstanding

During the preprocessing phase, the missing part always is omitted by researchers to make the process becomes easier. However, in this study, the data will become two sets of data which the first one is data with missing readings been deleted while the other data is the missing parts that were being replaced. The missing readings were replaced with mode for each attribute. It means that the most frequent data that appeared on each variable will be identified for substitution purposes. Mode values were used as the data consisted of nominal and categorical attributes. It means that mode values are the suitable values that can be obtained from the data. Therefore, the data will become completed by the process of replacing the missing part.

Meanwhile, the strings reading that represented the nominal value were changed into numerical data that are suitable for the attributes. As an example, the age of the patients was divided into 3 groups, whereby the age from 21 to 25 is replaced by 1, 26–30 is replaced by 2, and 31–35 is replaced by 3. The replacements made were shown in Results and Discussion.

25.2.2 Classification Process

The classification process of the data was performed using the Rstudio decision tree algorithm using rpart and tree algorithm to construct the decision tree. The classification tree, also known as a decision tree, contains the root node, decision node, or leaf nodes that contribute to decision making, which represents any class of the data.

Table 25.1 Confusion matrix

		Predicted class	
		Class = 1	Class = 0
Actual class	Class = 1	f_{11}	f_{10}
	Class = 0	f_{01}	f_{00}

For this analysis, the tree used the default setting from Rstudio, which applied the Classification and Regression Tree Algorithm (CART). For CART, the tree used the approach of the GINI Index in selecting the parameter to be placed as a child node. The attribute with the lower Gini Index was preferred. The GINI index can be calculated as follows:

$$\text{Gini} = 1 - \sum_{i=1}^c (p_i)^2 \tag{25.1}$$

where (p_i) is the probability of an object being classified to a particular class.

The accuracy of a tree can be calculated using the actual class and predicted class of the data. Thus, based on the tree produced, the confusion matrix was used. The confusion matrix helped compare which dataset has a higher result of correct classification. Table 25.1 shows the confusion matrix for the classification tree.

The accuracy can be calculated from the confusion matrix. The best accuracy score is 1 and the worst score is 0.

$$\text{Accuracy} = \frac{f_{11} + f_{00}}{f_{11} + f_{00} + f_{01} + f_{10}} \tag{25.2}$$

where

- f_{11} actual class and predicted class in Class = 1
- f_{00} actual class and predicted class in Class = 0
- f_{01} actual class is Class = 0 and predicted class is Class = 1
- f_{10} actual class is Class = 1 and predicted class is Class = 0.

Classification rules can be obtained from the decision trees that are in the form of if-then statements. Based on the decision rules, the associations between variables can be identified.

25.3 Results and Discussion

The breast cancer data were used for further analysis using Rstudio software with $n = 286$. The missing data that exist in the data were omitted because a decision tree cannot be built when there are missing values in the data. another way of handling the missing data was replacing the missing values with suitable values (Table 25.2).

It was found that the process of building the decision trees took a long time as all the variables used the string form that shows the nominal data. As a solution to this problem, the data used were transformed into a digit form. The changes made for every attribute were recorded in Table 25.3.

25.3.1 Decision Tree for Data with Missing Readings Detected

After the cleaning process, attributes with $n = 277$ were used for further analysis. The missing readings were deleted completely and no replacement was done. The process of building the decision tree was carried out and the prediction of breast cancer recurrence was identified.

Figure 25.1 shows the decision tree for breast cancer data that were classified into non-recurrence (NRC) and recurrence (RC) cancer. From the tree, the percentage for each split is shown clearly. The first slitting nodes showed that 70% of the data were classified into NRC while another 30% was in the RC class. However, 30% of the RC class can be further divided into a few other rules. The decision rules can be obtained from the decision tree. Based on the tree, only 4 attributes were used to build the decision tree from 9 attributes in the data. The unused attributes in the tree are age, breast, menopause, node caps, and irradiation, while the four attributes used are breast quad, degree malignant, involve nodes, and tumor size. This means

Table 25.2 Description of dataset

Attribute	Description
Age	9 groups of ages from 10 to 99 years old
Breast	Left or right breast
Breast quad	5 types of breast quads of both sides of breasts
Menopause	3 groups of menopause
Tumour size	12 groups of tumour size
Involve nodes	13 groups of involve nodes
Node caps	Existence of node caps
Degree of malignant	3 stages of malignant
Irradiation	Existence of irradiation

Table 25.3 Code for age based on group range

No	Attributes	Code
1	Age	"1": 10-19 "2": 20-29 "3": 30-39 "4": 40-49 "5": 50-59 "6": 60-69 "7": 70-79 "8": 80-89 "9": 90-99
2	Breast	"1": Left "2": Right
	Breast quad	"1": Left up "2": Left low "3": Right up "4": Right low
4	Menopause	"1": lt40 "2": ge40 "3": premeno
5	Tumor size (mm)	"1": 0-4 "2": 5-9 "3": 10-14 "4": 15-19 "5": 20-24 "6": 25-29 "7": 30-34 "8": 35-39 "9": 40-44 "10": 45-49 "11": 50-54 "12": 55-59
6	Involve nodes	"1": 0-2 "2": 3-5 "3": 6-8 "4": 9-11 "5": 12-14 "6": 15-17 "7": 18-20 "8": 21-23 "9": 24-26 "10": 27-29 "11": 30-32 "12": 33-35 "13": 36-38

(continued)

Table 25.3 (continued)

No	Attributes	Code
7	Node caps	“1”: Yes “2”: No
8	Degree of malignant	“1”: Low “2”: Moderate “3”: Severe
9	Irradiation	“1”: Yes “2”: No

that only 4 attributes are really important and contribute to identifying the class of the data that leads to non-recurrence or recurrence of breast cancer.

The decision rules from the trees are as follows:

1. If the degree of malignant is less than 3, then the breast cancer is classified as non-recurrence.
2. If the degree of malignant is more than 3, involve nodes more than 2, then the breast cancer is classified as recurrence.
3. If the degree of malignant is more than 3, involve nodes less than 2, tumor size less than 6, then the breast is classified as non-recurrence.
4. If the degree of malignant more than 3, involve nodes less than 2, tumor size more than 6, breast quad code is more than 3, then the breast is classified as recurrence.
5. If the degree of malignant is more than 3, involve nodes less than 2, tumor size more than 6, breast quad code is less than 3, then the breast is classified as recurrence.

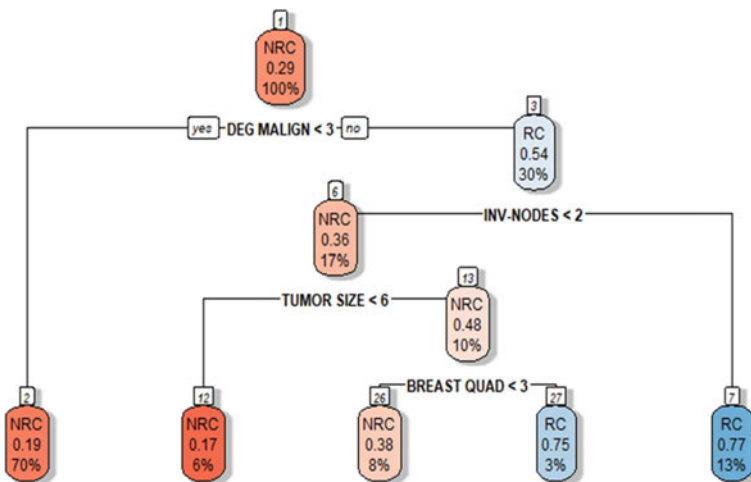


Fig. 25.1 Decision tree for breast cancer data with deleted missing readings

Table 25.4 Confusion matrix for decision tree accuracy

Class	Non-recurrence	Recurrence
Non-recurrence	20	2
Recurrence	2	3

Table 25.5 Complexity parameter table for data with missing readings

Complexity parameter	rel error	xerror	xstd
0.117284	1.00000	1.00000	0.093464
0.024691	0.76543	0.83951	0.088431
0.010000	0.71605	0.86420	0.089291

Root node error: $81/277 = 0.29242$

Table 25.4 shows the confusion matrix obtained from the testing process. 27 data, which denote 10% from the data, were used for testing. The confusion matrix shows the accuracy of the decision tree to classify the data. The percentage of accuracy obtained was 85.20%. From 27 data points used for testing, 23 data were correctly classified with 20 non-recurrences and 3 recurrences. The percentage can be considerably high in terms of its accuracy.

Table 25.5 shows the complexity parameter (CP) of the decision tree. The CP value can be used to control the tree growth by which pruning can be carried out based on the CP values whereby its corresponding cross-validated error (xerror) with the least value is considered as the optimal value of CP. For this study, the smallest value of xerror is 0.83951 by which the CP value is 0.024691 and *nsplit* is 2. Meanwhile, the root node error is 0.29242, which denotes that approximately up to 70% of the data were correctly classified in the first splitting node. This shows that the decision tree is capable of classifying breast cancer data.

25.3.2 Decision Tree with Mode for Missing Readings Replacement

In this section, another alternative was used to handle missing data. The modes for every variable were recognized and were used to replace every missing reading that existed in the data. Therefore, the number of data is the same as the original data obtained from the source which is $n = 286$. With the same CART algorithm, the analysis was carried out on the data.

Figure 25.2 shows the decision tree for breast cancer data with mode as a replacement for missing data. 70% of the data were classified as NRC at the first splitting branch while another 30% were RC. However, when the splitting nodes increased, the rules also changed. There are 5 variables shown in the decision tree which are degree malignant, involve nodes, breast, tumor size and breast quadrant. It means

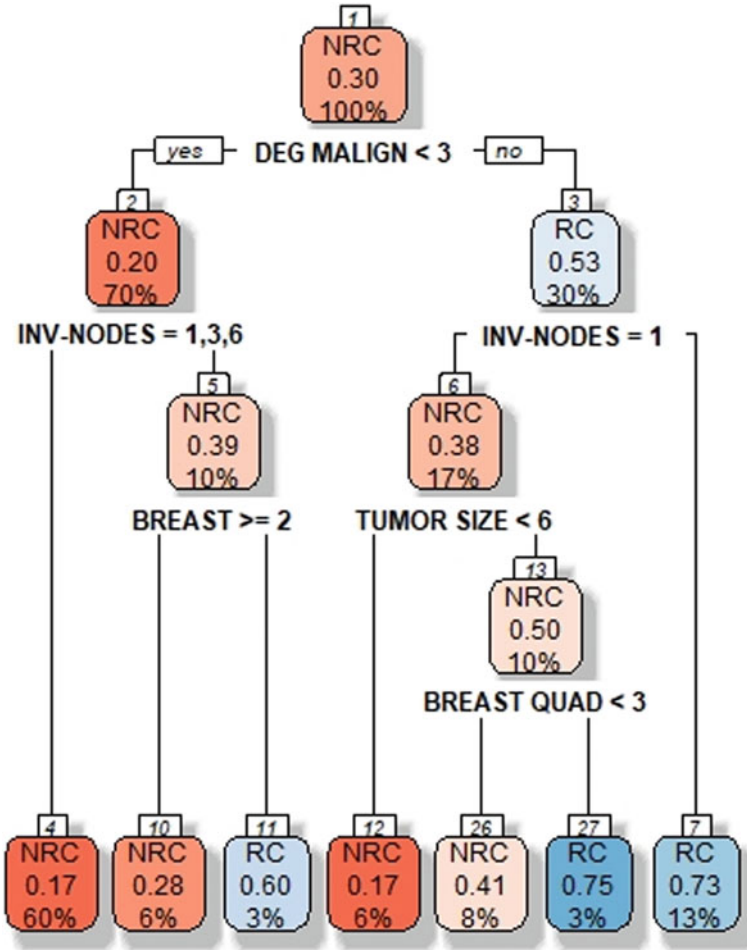


Fig. 25.2 Decision tree for breast cancer data with mode replacement for missing readings

that only these five variables are important in identifying the class of breast cancer data. The decision rules for the tree are as follows:

1. If the degree of malignant is less than 3, involve node is 1/3/6, then the breast cancer is classified as non-recurrence.
2. If the degree of malignant is less than 3, involve node is not 1/3/6, breast side is 2, then the breast cancer is classified as non-recurrence.
3. If the degree of malignant is less than 3, involve node is not 1/3/6, breast side is 1, then the breast cancer is classified as recurrence.
4. If the degree of malignant is less than 3, involve node is not 1, then the breast cancer is classified as recurrence.

5. If the degree of malignant is more than 3, involve nodes is 1, tumor size less than 6, then the breast is classified as non-recurrence.
6. If the degree of malignant more than 3, involve nodes less than 2, tumor size more than 6, breast quad code is more than 3, then the breast is classified as recurrence.
7. If the degree of malignant is more than 3, involve nodes is 1, tumor size more than 6, breast quad code is less than 3, then the breast is classified as non-recurrence.

Table 25.6 shows the confusion matrix obtained from the testing process. 28 data, which denote 10% from the data, were used for testing. The confusion matrix shows the accuracy of the decision tree to classify the data. The percentage of accuracy obtained was 78.57%. From 28 data points used for testing, 22 data were correctly classified with 18 non-recurrences and 4 recurrences. The percentage of accuracy can be considered as high and it indicates that the mode substitution can be used for categorical data if there is missing data that needs replacement.

Table 25.7 shows the complexity parameter (CP) of the decision tree when data with mode as a replacement is used. The CP value can be used to control the tree growth by which pruning can be carried out based on the CP values whereby its corresponding cross-validated error (xerror) with the least value is considered as the optimal value of CP. For this study, the smallest value of xerror is 0.90588 by which the CP value is 0.023529 and nsplit is 2. Meanwhile, the root node error is 0.2972, which denotes that approximately up to 70% of the data were correctly classified in the first slitting node. This shows that the decision tree is capable of classifying breast cancer data.

Table 25.6 Confusion matrix for decision tree accuracy for data with missing reading replacement

Class	Non-recurrence	Recurrence
Non-recurrence	18	5
Recurrence	1	4

Table 25.7 Complexity parameter table for data with missing readings replacement

Complexity parameter	nsplit	rel error	xerror	xstd
0.100000	0	1.00000	1.00000	0.090930
0.023529	2	0.80000	0.90588	0.088250
0.011765	4	0.75294	0.95294	0.089643
0.010000	6	0.72941	0.96471	0.089975

Root node error: $85/286 = 0.2972$

25.4 Conclusion and Recommendation

The breast cancer data containing 9 attributes with 2 class attributes were used in the decision tree algorithm. The preprocessing of the data was done to omit the missing reading that existed in the data and some modifications were made to the data to decrease the time taken to build the decision tree. The decision tree was visualised using the Rstudio software, thus showing that the non-recurrence and recurrence data can be classified. Only 4 attributes were displayed in the tree, which indicated that the other attributes that were not shown in the tree were less important in the classification process. The accuracy of the decision tree for the breast cancer data is more than 0.5, which is rather good for the data. By using the decision tree, the list of classification rules can be identified because, in the future, the rules can easily be used to identify whether the patients' breast cancers might be non-recurrence or recurrence cancer. Besides, the root node error is only 0.29242, hence showing that the data were highly and correctly classified from the first splitting node.

Meanwhile, the same data with the substitution technique using mode for each variable has also been used and gained another result of the decision tree. Only 5 variables were displayed in the decision tree which shows that only 5 variables are significant in building the tree. The smallest error obtained from the second decision tree is 0.90588 when the $n_{split} = 2$. The smallest relative error is 0.72941 with $n_{split} = 6$. The root node error for the second tree is 0.2972 which is slightly higher than the first decision tree. The data correctly classified was also 70% which can be considered as high accuracy.

From the first decision tree, it can be concluded that if the breast cancer is a non-recurrence, the degree of malignant must be less than 3 which is low and moderate. Meanwhile, when the degree of malignant is high and involves nodes more than 2, the breast cancer could be a recurrence. Therefore, the future patients that have these criteria might have the same possibilities either the cancer is recurrence or non-recurrence.

For future research, it is recommended to use bigger size data that might affect the classification process. In this study, two ways of handling data with missing values were used before further analysis which was missing values were omitted without considering any alternatives to replace the reading and replaced the missing reading with the mode of respective variable as the second alternative. Both of the decision trees gave a good result and the breast cancer data managed to be classified. The decision rules used also can be clearly recognized. Therefore, it is interesting to explore the result of the decision tree if the missing reading is replaced with some values using other methods. As this study used online-sourced data, it would be very fruitful if this kind of study is carried out using Malaysia's breast cancer data.

Acknowledgements The authors would like to thank the Research Management Centre UTHM (RMC) for giving the chance to carry out this research by using Geran Penyelidikan Pascasiswazah (GPPS). The authors would also like to thank UCI Machine Learning for becoming a source for the researchers to explore more useful data in the respective field. The authors profusely thank the University Medical Centre, Institute of Oncology, Ljubljana, Yugoslavia for providing this data as

an open-source to the authors. The authors are also extremely thankful to the reviewers for their beautiful remarks.


References

1. D. Hand, H. Manilla, P.Smyth, *Principles of Data Mining* (The MIT Press, England, 2001)
2. J. Han, M. Kamber, *Data Mining: Concepts and Techniques* (Morgan Kaufmann Publishers, USA, 2001)
3. L. Adhatrao, A. Gaykar, A. Dhawan, R. Jha, V. Honrao. Predicting students' performance using ID3 and C4.5 classifications algorithms. *Int. J. Data Mining Knowl. Manage. Process (IJDKP)* **3**(5), 39–52 (2013)
4. D.T. Larose, *Data Mining Methods and Models* (Wiley, The United States of America, 2006)
5. E. Karoussi, A. Bouhmala, Data mining K-clustering problem. Master's Thesis, University of Agder (2012)
6. S.C. Suh, *Practical Applications of Data Mining*. USA. Jones & Bartlett Learning (2012)
7. J. Solarte, A proposed data mining methodology and its application to its industrial engineering. Master's Thesis, University of Tennessee (2002)
8. Two-Crows, *Introduction to Data Mining and Knowledge Discovery* (Two Crows Corporation, USA, 1999)
9. S. Sharma, J. Agrawal, S. Sharma, Classification through Machine Learning Technique: C4.5 (2013)
10. I.A. Kareem, M.G. Duaimi, Improved accuracy for decision tree algorithm based on unsupervised discretization. *Int. J. Comput. Sci. Mob. Comput.* **3**(16), 176–183 (2014)
11. M. Kantardzic, *Data Mining: Concepts, Models, Methods, and Algorithms* (Wiley, New York, 2003)
12. E.M. Karabulut, T. Ibryikei, Analysis of cardiocogram data for fetal distress determination by decision tree based adaptive boosting approach. *J. Comput. Commun.* **2**, 32–37 (2014)
13. V. Chaurasia, S. Pal, Data mining approach to detect heart diseases. *Int. J. Adv. Comput. Sci. Inf. Technol. (IJACSIT)* **2**(4), 56–66 (2013)
14. A. Saygili, Classification and diagnostic prediction of breast cancers via different classifiers. *Int. Sci. Vocat. J. (ISVOS)* **2**(2), 48–56 (2018)
15. A.K.S. Banu, S.H. Ganesh, A hybrid approach for an efficient classification using decision tree and SVM. *Int. J. Comput. Sci.* **7**(2), 42–48 (2018)
16. L. Breiman, *Classification and Regression Trees*. Routledge (2017). .
17. S. Singh, P. Gupta, Comparative study ID3, CART and C4.5 decision tree algorithm: a survey. *Int. J. Adv. Inf. Sci. Technol. (IJAIST)* **27**(27), 97–103 (2014)
18. G. Andrew, H. Jennifer, Missing data imputation. data analysis using regression and multilevel/hierarchical models, 529–544 (2010)
19. I.F. Ilyas, X. Chu, Trends in cleaning relational data: consistency and deduplication. *Found. Trend Data Bases* **5**(4), 281–393 (2015)
20. J.W. Osborne, *Best Practices in Data Cleaning* (SAGE Publications, USA, 2013)

Chapter 26

Socio Demographic Determinants of Contraceptive Use Among Women in Malaysia: Logistic Regression Model



Nor Azah Samat, Noor Azlin Muhammad Sapri, Norziha Che Him , Adzmel Mahmud, and Noor Wahida Md. Junus

Abstract Family planning could be defined as the ability of individuals and couples to anticipate and achieve their desired number of children at the same time spacing and timing of their births. There are many types of contraceptive methods that can be used in family planning that could divided into two categories; namely modern and non-modern contraceptive methods. The choice whether to use modern or non-modern methods is said to be related to the socio demographic characteristics of the woman. This study aims to investigate the socio-demographic determinants of contraceptive use among ever married women aged 15–49 in Malaysia. Data used in this study was obtained from a national survey called Fifth Malaysia Population and Family Survey (MPFS-5) which was conducted by the National Population and Family Development Board in 2014. Multi-stage stratified random sampling method was used across all states in Malaysia. A total of 2736 ever married women who use contraception aged 15–49 years old were subjected to a face-to-face interview. The information was obtained on socio demographic characteristics and family planning practice. Data was analysed by using descriptive analysis, Chi-Square analysis and binary logistic regression. Findings of the analysis revealed that amongst five socio-demographic

N. A. Samat (✉) · N. W. Md. Junus
Department of Mathematics, Faculty of Science and Mathematics, Universiti Pendidikan Sultan Idris, 35900 Tanjong Malim, Perak, Malaysia
e-mail: norazah@fsmt.upsi.edu.my

N. W. Md. Junus
e-mail: noor_wahida@fsmt.upsi.edu.my

N. A. M. Sapri · A. Mahmud
Population and Family Research Division, National Population and Family Development Board, 50350 Kuala Lumpur, Malaysia
e-mail: azlin@lppkn.gov.my

A. Mahmud
e-mail: adzmel@lppkn.gov.my

N. C. Him
Department of Mathematics and Statistics, Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, 84600 Pagoh, Johor, Malaysia
e-mail: norziha@uthm.edu.my

characteristics under study which are stratum, ethnicity, education level, age and income, only three socio-demographic characteristics known as stratum, ethnicity and education level present a significant association to the modern and non-modern contraceptive use among ever married women in Malaysia. In conclusion, the stratum or the location which either from urban or rural areas, the ethnicity which either Malay, other Bumiputera, Chinese or Indian, and the education level which either no education, primary school, secondary school or tertiary school, are the determinants of modern and non-modern contraceptive use among ever married women aged 15–49 in Malaysia.

26.1 Introduction

Family planning is important in ensuring the health of the mother is guaranteed and the life of the child is not neglected, especially for women who are married. There are various types of family planning methods provided in most clinics in Malaysia, especially the Nur Sejahtera Clinic under the National Population and Family Development Board Malaysia (NPFDB). The choice of contraceptive methods use will normally be based on doctor's advice and the woman's choice. The types of contraceptive methods can be divided into two categories; namely modern methods and non-modern methods. Modern contraceptive methods are often defined as a technological advancements that aim to transcend biology in which its enable couples to have sexual intercourse at any mutually-desired time, with diminished risks of pregnancy [1]. Conversely, non-modern methods are defined as a natural contraceptive methods, which is more to traditional methods. Among the modern methods that are commonly used by most women are hormone injection, condoms, pills, rhythm and many more [2, 3], whereas the commonest non-modern methods include traditional methods such as massaging, the use of herbs, and so on [4, 5].

The choice of contraceptive or family planning method may be caused by certain characteristics or factors such as socio demographic, economy factor and others. There are many studies have been done on determinants of contraceptive use among women in many countries including Zambia, Malaysia, India, USA, and many more [6–9]. In Malaysia for instance, study by [10] shows that marital status, religion, number of children and attitude were identified as significant factors of family planning practice. While, study in Uganda by [11] found out that the education, prior use of contraceptive and partner communication are the determinants of family planning among postpartum women. Conversely, instead of women, there is also a study on the factors associated with contraceptive use among married men visiting a healthcare clinic in Malaysia [12]. Study in determining the determinant use of contraceptive is important to assist authorities especially family planning clinic in providing better family planning services [13, 14]. Therefore, this study aims to identify the determinants of contraceptive use among married women in Malaysia.

26.2 Methodology

The methodology used in this study includes descriptive analysis of the demographic characteristics of respondents, bivariate correlations and binary logistic regression analyses [15]. Meanwhile, the data used were secondary data obtained from the latest wave of Fifth Malaysia Population and Family Survey (MPFS-5), a nationally representative survey conducted by the National Population and Family Development Board in 2014. The sample for this survey was selected by the Department of Statistics Malaysia (DOSM), in which a multi-stage stratified random sampling was adopted to select the living quarters in the MPFS-5 and all Malaysian ever married women aged 15–49 years old in the selected household were interviewed. The survey was conducted through face-to-face interviews by trained interviewers using standardized questionnaires in all states in Malaysia. Data for a total of 5558 respondents were extracted from the ever married women sample of the survey. Then, these respondents were filtered to include only 2736 women who use contraception, and is the main focus in this study.

A logistic regression analysis could be used to examine the association of categorical or continuous independent variables that consist of one dichotomous dependent variable (binary response). Here, the logistic regression model compares the odds of a prospective attempt in those with and without prior attempts with the ratio of those odds known as the odds ratio. It does not analyse the odds but a natural logarithmic transformation of the odds or the log odds [16]. Since it applied the logarithmic transformation of the odds ratio, therefore the interpretation of results from the computer output needs a transformation back to the original scale by taking the inverse of the natural log of the regression coefficient. This step known as exponentiation where the exponentiated regression coefficient could represents the strength of the association of the independent variable with the response in terms of increase or decrease in risk of the outcome that is associated with the independent variable. The main advantage is to avoid confounding effects by analysing the association of all variables together [17]. Several link functions are available, but the most common link for binary is the logit link function, defined as in Eq. 26.1.

$$g(\mu_i) = \text{logit}(\mu_i) = \log\left(\frac{\mu_i}{1 - \mu_i}\right) \quad (26.1)$$

$$\log\left(\frac{\mu_i}{1 - \mu_i}\right) = \alpha + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p \quad (26.2)$$

The logistic regression equation in logit form as shown in Eq. 26.2, where, α is an intercept parameter, β is a regression coefficient, μ_i is the probability of an event occurring and $1 - \mu_i$ is the probability of an event not occurring.

The dependent variable is types of contraceptive method use among women aged 15–49 in Malaysia. In the MPFS5 questionnaire, there are 19 types of contraceptive methods listed which include abstinence, condom, pill, emergency contraceptive pill,

implant, hormone injection, intra uterine device (IUD), contraceptive cream, contraceptive jelly, hormone patch, vaginal ring, tube ligation, vasectomy, contraceptive foam, rhythm, withdrawal, breastfeeding, traditional and others. In this study, these types of contraceptive methods have been categorized into two categories which are modern method and non-modern method (Y_i).

Table 26.1 presents the list of variables and category for this study where the independent variables (X_i) that considered in this study are the socio-demographic

Table 26.1 List of variables and category for the study

No	Variables	Category
1	Number of client according to types of contraceptive use 2 categories of contraceptive method: 1—modern method 2—Non-modern method	1—Modern method Abstinence Condom Pill Emergency contraceptive pill Implant Injection Intrauterine device (IUD) Contraceptive cream Contraceptive jelly Hormone patch Vaginal ring Tube ligation Vasectomy Contraceptive foam Rhythm Withdrawal Breastfeeding 2—Non-modern method Traditional Other
2	Stratum	Urban Rural
3	Ethnicity	Malay (Bumiputera) Other Bumiputera Chinese Indian
4	Age	15–19 20–29 30–39 40–49
5	Education Level	No education Primary school Secondary school Tertiary
6	Income	B40: RM4360 and below M40: RM4361–RM9619 T20: RM9620 and above

characteristics of ever married women who use the contraception which includes five variables known as stratum, ethnicity, age, education level, and income. The data analyses on these five characteristics are next analysed to investigate the preference use of modern and non-modern contraceptive methods in Malaysia.

26.3 Data Analysis and Findings

Table 26.2 shows the percentage of respondents according to socio-demographic characteristics based on types of contraceptive use for ever married women aged 15–49 years old who use contraception. According to the stratum, 61.9% respondents are from urban area, while 38.1% are from rural area. Among these women, 67.5% use modern contraceptive methods compared to non-modern contraceptive methods.

Furthermore, majority of respondents (63.7%) are Malay women, followed by other Bumiputera (18.5%), Chinese (11.8%) and Indian (6.0%). Among these Malay women, 41.4% use modern contraceptive and 22.3% use non-modern contraceptive. In addition, most women involve in this study are among women aged 30–39 years

Table 26.2 Distribution of percentage for respondents based on socio-demographic characteristics and types of contraceptive use

	Categories	(Percentage)	
Variable	Modern method	Non-modern method	Total
Stratum	Urban (39.9%) Rural (27.6%)	Urban (22.0%) Rural (10.5%)	Urban (61.9%) Rural (38.1%)
Ethnicity	Malay (41.4%) Other Bumiputera (14.4%) Chinese (8.6%) India (3.1%)	Malay (22.3%) Other Bumiputera (4.1%) Chinese (3.3%) India (2.9%)	Malay (63.7%) Other Bumiputera (18.5%) Chinese (11.8%) India (6.0%)
Age	15–19 (0.4%) 20–29 (12.7%) 30–39 (27.6%) 40–49 (26.8%)	15–19 (0.1%) 20–29 (5.1%) 30–39 (13.2%) 40–49 (14.1%)	15–19 (0.5%) 20–29 (17.8%) 30–39 (40.8%) 40–49 (40.9%)
Education	No education (1.6%) Primary school (8.0%) Secondary school (43.8%) Tertiary school (14.2%)	No education (0.5%) Primary school (27.0%) Secondary school (19.7%) Tertiary school (9.5%)	No education (2.1%) Primary school (10.7%) Secondary school (63.5%) Tertiary school (23.7%)
Income	B40 (24.4%) M40 (2.4%) T20 (40.7%)	B40 (13.3%) M40 (1.8%) T20 (17.3%)	B40 (37.8%) M40 (4.2%) T20 (58.0%)

Table 26.3 Chi-square test summary on association socio-demographic characteristics and types of contraceptive use (modern and non-modern)

Variables	Chi-square	Significance
(1) Stratum	18.759	0.000*
(2) Ethnicity	50.438	0.000*
(3) Age	6.898	0.075
(4) Education level	26.116	0.000*
(5) Income	14.234	0.001*

*Significance at $\alpha = 0.05$

old (40.8%) and 40–49 years old (40.9%). In terms of education, majority education level of the women is from secondary school (63.5%), in which most of them use modern contraceptive method (43.8%). Meanwhile, in terms of level of income, women participated in this study was working with household income RM9620 and above which is category T20 (58.0%), followed by category B40 (37.8%) and M40 (4.2%).

Next, Table 26.3 presents the results of analysis by Chi-square test to determine whether the independent variables involved are significantly associated to the dependent variable which is whether or not the types of contraceptive method used by women are associate with the socio-demographic characteristics. Findings show that the socio-demographic characteristics such as stratum ($p > 0.000$), ethnicity ($p > 0.000$), education level ($p > 0.000$) and income ($p > 0.001$) have significant association on the types of contraceptive use; modern and non-modern methods among ever married women aged 15–49 in Malaysia, except for age ($p > 0.075$). Therefore, variable age is dropped and will not be analysed in the binary logistic regression.

Next, further analysis on this association are next discussed by using binary logistic regression to determine the significant determinants of those four characteristics in the model. Table 26.4 shows the results of analysis using Full Model of Logistic Regression. The findings obtained indicated that several variables insignificant and should be dropped from the full model. The variable which is found not statistically significant is household income, meanwhile socio demographic characteristics consists of stratum ($p > 0.003$), ethnicity ($p > 0.000$) and education level ($p > 0.048$) are found to be significantly associated to the contraceptive use and will be further analyses as a Final Model of Logistic Regression (see Table 26.5).

Thus, results on the best fitted model, we called here as the final model, involving the important variables are summarised in Table 26.5. It can be seen that three of four of socio-demographic characteristics which are stratum, ethnicity and education level have significant association to the contraceptive use; modern or non-modern methods, among ever married women in Malaysia. Obviously, women from rural area are said to be 1.319 times more likely to choose modern methods compared to women from urban area.

Furthermore, Other Bumiputera women are 2.777 times more likely to choose modern methods, followed by Chinese with 2.486 times, and Malay with 1.627 times, compared to Indian women. While, women with primary school education is 1.034 more likely to use modern contraceptive method compared to women without

Table 26.4 A full model of logistic regression of socio demographic characteristics for women with contraception

Covariates category	Significance	β Model	Standard error (S.E.)
Intercept		0.024	0.403
Stratum	0.003*		
– Urban (RC)	–	–	–
– Rural	0.003*	0.272	0.091
Ethnicity	0.000*		
– Malay	0.003	0.491	0.167
– Other Bumiputera	0.000	1.021	0.193
– Chinese	0.000	0.913	0.202
– India (RC)	–	–	–
Education level	0.048*		
– No education (RC)	–	–	–
– Primary school	0.968	0.014	0.343
– Secondary school	0.711	– 0.120	0.323
– Tertiary	0.256	–0.378	0.333
Household income	0.122		
– B40	0.663	0.091	0.209
– M40 (RC)	–	–	–
– T20	0.223	0.259	0.212

Notes RC = Reference category
 *Significance at $\alpha = 0.05$

education, and women with secondary and tertiary school education are 0.895 and 0.648 times less likely to use modern methods compared to non-modern methods, subsequently.

Therefore, the final logistic regression model is as the following,

$$\begin{aligned}
 \log\left(\frac{\mu_i}{1 - \mu_i}\right) &= 0.214 + 0.277(\text{Stratum Rural}) + 0.487(\text{Ethnicity Malay}) \\
 &+ 1.021(\text{Ethnicity Other Bumiputera}) \\
 &+ 0.911(\text{Ethnicity Chinese}) \\
 &+ 0.034(\text{Education Primary School}) \\
 &- 0.111(\text{Education Secondary School}) \\
 &- 0.435(\text{Education Tertiary})
 \end{aligned}
 \tag{26.3}$$

Table 26.5 A final model of logistic regression of socio demographic characteristics for women with contraception

Covariates category	Coefficients	Significance	<i>B</i>	Exp(<i>B</i>)	Standard error (S.E.)
Intercept	α		0.214	1.239	0.346
Stratum		0.002*			
– Urban (RC)		–	–	–	
– Rural	β_{1Rural}	0.002*	0.277*	1.319	0.091
Ethnicity		0.000*			
– Malay	β_{2Malay}	0.004*	0.487*	1.627	0.167
– Other Bumiputera	$\beta_{2Other\ Bumiputera}$	0.000*	1.021*	2.777	0.193
– Chinese	$\beta_{2Chinese}$	0.000*	0.911*	2.486	0.201
– India (RC)		–	–	–	
Education level		0.003*			
– No education (RC)		–	–	–	
– Primary school	$\beta_{3Primary\ School}$	0.922	0.034	1.034	0.343
– Secondary school	$\beta_{3Secondary\ School}$	0.730	–0.111	0.895	0.323
– Tertiary	$\beta_{3Tertiary}$	0.189	–0.435	0.648	0.331

Notes RC = Reference Category

*Significance at $\alpha = 0.05$

26.4 Conclusion and Discussion

Women in Malaysia have the option to choose whether they want to use modern or non-modern contraceptive methods for their family planning, with doctor advice. Finding of the study show that choosing modern or non-modern contraceptive methods among ever married women aged 15–49 years old depends on whether women are either from urban or rural area. Women from rural area are said to be more likely to use modern method compared to women from urban area. While, other Bumiputera women are said to be more likely to use modern method followed by Chinese and Malay compared to Indian women. Furthermore, women with primary school education is more likely to use modern contraceptive method compared to women without education, and women with secondary and tertiary school education less likely to use modern methods compared to non-modern methods. This results are important and could be used by the authorities or family planning clinic to provide excellent services towards better family well-being. For instance, the findings of this study related to the geographical location, ethnicity and levels of education towards the types of contraception methods either modern or non-modern can be used as input in improving the services in terms of allocation of resources, training, and awareness campaigns. However, this study only consider several important socio-demographics characteristics. Future research may consider other variables such as

religion, number of children and prior use of contraceptive, so that the study can be more beneficial.

Acknowledgements The authors acknowledge Universiti Pendidikan Sultan Idris for their financial support in respect of this study.

References

1. H. David, T. James, A definition of modern contraceptive methods. *Contraception* **92**(5), 420–421(2015)
2. D. Hubacher, J. Trussell, A definition of modern contraceptive methods. *Contraception* **92**, 420–421 (2015)
3. N. Cahill, E. Sonneveldt, J. Stover, M. Weinberger, J. Williamson, C. Wei, W. Brown, L. Alkema, Modern contraceptive use, unmet need, and demand satisfied among women of reproductive age who are married or in a union in the focus countries of the family planning 2020 initiative: a systematic analysis using the family planning estimation tool. *Lancet* **391**, 870–882 (2018)
4. A.I. Ajayi, O.V. Adeniyi, W. Akpan, Use of traditional and modern contraceptives among childbearing women: findings from a mixed methods study in two Southwestern Nigerian States. *BMC Public Health* **18**, 604–612 (2018)
5. N. Cohen, F.T. Mendy, J. Wesson, A. Protti, C. Cissé, E.B. Gueye, L. Trupe, R. Floreak, D. Guichon, K. Lorenzana, A. Buttenheim, Behavioral barriers to the use of modern methods of contraception among unmarried youth and adolescents in Eastern Senegal: a qualitative study. *BMC Public Health* **20**, 1025–1033 (2020)
6. D.A. Jawa, M.M. Rahman, Factors affecting contraceptive use among the women of reproductive age in Samarahan District, Sarawak, Malaysia. *Malaysian J. Public Health Med.* **15**(1), 10–17 (2015)
7. J. Lasong, Y. Zhang, S.A. Gebremedhin, S. Opoku, C.S. Abaidoo, T. Mkandawire, K. Zhao, H. Zhang, Determinants of modern contraceptive use among married women of reproductive age: a cross-sectional study in rural Zambia. *BMJ Open* **10**, 1–10 (2020)
8. M.A. Ayanore, P.A. Aryee, Determinants and use of family planning among young women (18–28 years) attending health facilities in Garu-Tempne District of Ghana. *Int. J. Health Sci.* **3**(4), 43–54 (2015)
9. S.M. Valecha, M. Ahmed, A. Sood, Study of socio-economic determinants in family planning. *Int. J. Commun. Med. Public Health* **2**(2), 107–111 (2015)
10. M. Mansor, K.L. Abdullah, S.S. Oo, K. Akhtar, A.S. Jusoh, S. Ghazali, M. Haque, L.C. Choon, The prevalence of family planning practice and associated factors among women in Serdang, Selangor. *Malaysian J. Public Health Med.* **15**(3), 147–156 (2015)
11. K.M. Sileo, R.K. Wanyenze, H. Lule, S.M. Kiene, Determinants of family planning service uptake and use of contraceptives among postpartum women in rural Uganda. *Int. J. Public Health* **60**(8), 987–997 (2015)
12. M.H. Syahnaz, J.K. Rasina Nilofer, M.W. Azmawati, H.H. Siraj, Factors associated with ever used of modern contraception among married men attending a primary healthcare clinic. *Med. J. Malaysia* **73**(5), 301–306 (2018)
13. A.M. Sidibe, P.I. Kadetz, T. Hesketh, Factors impacting family planning use in Mali and Senegal. *Int. J. Environ. Res. Public Health* **17**(12), 4399–4414 (2020)
14. M. Alemayehu, H. Lemma, K. Abrha, Y. Adama, G. Fisseha, H. Yebyo, E. Gebeye, K. Negash, J. Yousuf, T. Fantu, T. Gebregzabher, A.A. Medhanyie, Family planning use and associated factors among pastoralist community of Afar Region, Eastern Ethiopia. *BMC Womens Health* **16**(39), 1–9 (2016)

15. M.G. Anasel, U.J. Mlinga, Determinants of contraceptive use among married women in Tanzania: policy implication. *Afr. Popul. Stud.* **28**(2), 978–988 (2014)
16. S.M. Hailpern, P.F. Visintainer, Odds ratios and logistic regression: further examples of their use and interpretation. *Stata J.* **3**(3), 213–225 (2003)
17. S. Sperandei, Understanding logistic regression analysis. *Biochemia Medica* **24**(1), 12–18 (2014)

Chapter 27

Crude Palm Oil Price Forecasting: Comparative Study of Hybrid GMDH Neural Network and ARIMA Model



Huma Basheer  and Azme Khamis 

Abstract This paper attempts to develop a more appropriate hybrid wavelet modified GMDH model and its comparative study with statistical ARIMA model for the Malaysia monthly crude palm oil (CPO) price forecasting. The complex data of monthly CPO price has decomposed by using discrete wavelet analysis into different sub series in such a way that error criteria is minimized. The new input of monthly CPO price data obtained from the wavelet analysis is linked to the modified GMDH model. The importance of modified GMDH neural network that has been used different transfer functions simultaneously into GMDH and has given the best fit for modeling in comparison with conventional GMDH model. Real time series data of monthly CPO price for the period of 37 years are collected from Malaysian Palm Oil Board (MPOB). The performance of hybrid GMDH type neural network, individual GMDH model and statistical ARIMA model are determined by using mean absolute error (MAE), root mean squared error (RMSE), mean absolute percentage error (MAPE), correlation coefficient (R) and coefficient of determination (R^2). The result of hybrid GMDH neural network have shown an improvement over ARIMA model with 25.35%, 19.16% and 22.50% reduction in MAE, RMSE and MAPE values respectively. The obtained results of the comparison show that the hybrid wavelet-modified GMDH model has performed better result than the ARIMA model and individual GMDH neural network. The hybrid wavelet modified GMDH neural network technique is imperative to the field of forecasting as it can be used in different forecasting applications.

H. Basheer · A. Khamis (✉)

Department of Mathematics and Statistics, Faculty of Applied Sciences and Technology,
Universiti Tun Hussein Onn Malaysia (UTHM) Edu Hub Pagoh, 84600 Pagoh, Johor, Malaysia
e-mail: azme@uthm.edu.my

H. Basheer

e-mail: hw140007@siswa.uthm.edu.my

27.1 Introduction

Palm oil is one of the leading and most important vegetable oil traded in the worldwide market [1, 2]. Palm oil is a very productive crop and used in large amount as compared to any other vegetable oil due to its greater yield at a low cost of production. Global production and demand for palm oil is increasing rapidly. Malaysia has played and still considered the world palm oil leader due to the huge amount of production and export quantity [3–7]. In agriculture sector, Malaysia is gained largest income from the palm oil products and has played a major role in helping the economic growth of the country [8, 9]. Being a largest revenue of the agricultural sector in Malaysia, palm oil has an imperative role in industrial activities [10]. Malaysian government has taken the advantage to increase the role of palm oil industry in the world global market to enhance the economic growth of the country. Besides a major source of income, it provides major employment opportunities to Malaysian people, whose earning depends on palm oil industry.

Accurate and precise forecasting of palm oil price is important for enhancing the Malaysia economy [11]. The suitability of the forecasting methods and their modelling in predicting the palm oil price are becoming an important issue due to the price fluctuation in the future, which affect the international market all over the world. Conventionally for palm oil price forecasting different statistical models based on time series analysis has been used. These statistical models include autoregressive (AR), autoregressive moving average (ARMA), autoregressive integrated moving average (ARIMA), broken line (BL), fractional Gaussian noise (FGN), transfer function noise (TFN) and autoregressive moving average with exogenous terms (ARMAX) [12–15]. Since most of the current statistical models have been restricted to a linear models due to the limited effectiveness to forecast the CPO price, which is vastly nonlinear and also non-stationary. The most extensive statistical model, which is widely used for and prominent among all the class of linear model is the Box-Jenkins ARIMA model [16–18]. But ARIMA model due to limited effectiveness only taking the linear aspect of given data to forecast.

In the recent years, group method of data handling (GMDH) techniques have been successfully developed for modeling of non-linear time series forecasting. GMDH algorithm is used for time series forecasting in a vast range like signal processing, economy, control systems, medical diagnostics and ecology [19–23]. Parks et al. [24] and Ikeda et al. [25] have been studied and used the GMDH to find a model for the British economy and river flow prediction respectively. Similarly for prediction the currency of foreign exchange rate, Robinson and Mort [26] have been used an optimized GMDH algorithm to forecast the data for time series analysis. Fuzzy GMDH algorithm has been used by Hayashi et al. [27] to predict the manufacturing of computers in Japan. However, conventional GMDH model has some limitation such as it produces complex polynomial for simple experimental data and for high order nonlinear system it produce complex network model. In order to assist the problems associated with the conventional GMDH, Kondo et al. [28] suggested transfer functions like radial basis function, sigmoid function, instead of a single polynomial

to minimize the prediction error criterion. However, this type of modification is still in progress for different transfer function using non-stationary data. Currently, there are several reviews about the hybrid modeling, which suggest that hybrid system obtained better performance as compared to the traditional or conventional system [29, 30].

In this study, we proposed a hybrid wavelet modified GMDH model and its comparative study with ARIMA model for Malaysia monthly CPO price forecasting. Addressing the problems with the conventional GMDH based on quadratic polynomial, we proposed modified GMDH model based on the transfer functions (radial basis, sigmoid, polynomial and tangent) in order to improve the prediction accuracy of conventional GMDH model. The complex data of monthly CPO price has decomposed by using discrete wavelet analysis into different sub series in such a way that error criteria is minimized. The new input of monthly CPO price data obtained from the discrete wavelet analysis is linked to the modified GMDH model. The importance of modified GMDH neural network that has been used different transfer functions simultaneously into GMDH and has given the best fit for modeling in comparison with conventional GMDH model. The statistical performance measures such as RMSE, MAE, R, R^2 and MAPE are used to evaluate the hybrid GMDH neural network, individual GMDH model and statistical ARIMA model for the Malaysia monthly CPO price forecasting. The proposed hybrid GMDH neural network is expected to show more effective and efficient way to improve the forecasting accuracy and will be suitable for practical applications.

27.2 Methodology

For comparative study of hybrid GMDH neural network and ARIMA model, first we briefly discuss individual GMDH, modification of GMDH model, discrete wavelet transform (DWT) technique and hybridization of discrete wavelet transform with modified GMDH model. The statistical Box-Jenkins ARIMA (Autoregressive Integrated Moving Average) model is used for comparison purpose for monthly CPO price forecasting.

27.2.1 Group Method of Data Handling and Its Modification

GMDH is a mathematical modeling method that presents a useful approach to the recognition of high order non-linear system. GMDH is a multilayer network which has been successfully developed for modeling of linear and non-linear time series forecasting [31]. The distinguishing feature of group method of data handling is a heuristic self-organization method which implements an automatic selection of necessary input variables without using any event in advance of any empirical evidence, often

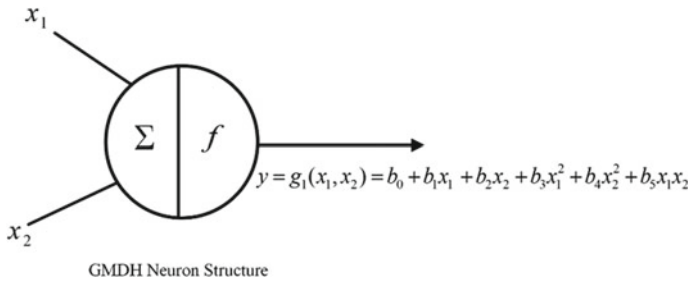


Fig. 27.1 GMDH neuron structure

subjectively among the input and output variables [32–34]. Its intention to make relation between regularly possible set of inputs and one output. The GMDH network basically determines the feasible inputs for the relevant system. During the training of the time series data, the GMDH network is developed layer by layer for the best possible output. Each individual layer of GMDH network consists of many neurons and each neuron has two inputs. The output of each neuron is a quadratic function of its both inputs as shown in Fig. 27.1 and calculated by Ivakhnenko polynomial as given in Eq. 27.1 [35].

$$y = b_0 + b_1x_i + b_2x_j + b_3x_ix_j + b_4x_i^2 + b_5x_j^2 \tag{27.1}$$

Here, y is a response variable, x_i and x_j are the covariate variables to be regressed. The Coefficients are $\{b_0, b_1, b_2, b_3, b_4, b_5\}$, which are determined by least square method. The response variable (y) is modelled by all possible combination of two input variables (x_i and x_j). It means that two input variables go in a neuron, one result goes out as an output.

The GMDH network is like neural network, where most of the units interchange their total inputs by utilizing a scalar to scalar function, which is known as transfer or activation function. The purpose of these net inputs interchangeability to pick up the best and correct model to overcome the non-linearity of the data, which clarify the relationship between the given inputs and proper outputs. By using the same procedure of GMDH neural network, modified GMDH model based on the transfer functions (radial basis, sigmoid, polynomial and tangent) in order to improve the prediction accuracy of conventional GMDH model [36]. The advantages of using these simultaneous transfer functions are also to overcome multi-collinearity problem using integrated regularized least square estimation and produce more precise and stable prediction for crude palm oil price forecasting as compared to conventional GMDH network [37]. The modified GMDH neural network is obtained by using four transfer functions simultaneously, where each neuron consists of four models and selects one model of that transfer function, which has the smallest external criteria. The selected transfer functions, which are used simultaneously as model input are shown in Table 27.1.

Table 27.1 Transfer functions

Transfer function	$z = f(y)$	Transformations
Sigmoid function	$z = \frac{1}{1+e^{-y}}$	$y = \ln\left(\frac{z}{1-z}\right), z \neq 1$
Radial basis function	$z = e^{-y^2}$	$y = \sqrt{-\ln(z)}$
Tangent function	$z = \tan(y)$	$y = \tan^{-1}(z)$
Polynomial	$z = y$	$y = z$

27.2.2 Discrete Wavelet Analysis

The main steps and study framework used in the discrete wavelet analysis is shown in Fig. 27.2. The discrete wavelet transform (DWT) is an application of the wavelet transform, in which the input time series data of monthly CPO price are decomposed into different sub series in such a way that error criteria is minimized. The original data is divided into two components after the decomposition process known

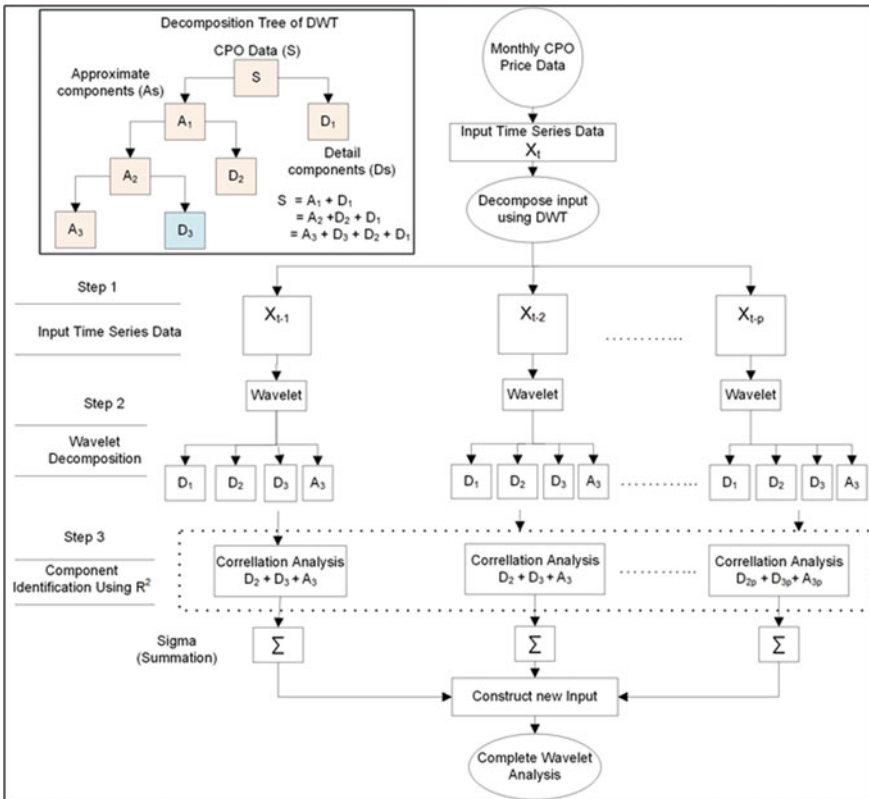


Fig. 27.2 Discrete wavelet analysis

as approximate (As) and detail coefficients (Ds). The effective components of the wavelet are determined using correlation coefficient (R) or coefficient of determination (R^2). The new input is constructed after the summation of most effective components of wavelet analysis, which are used as a clean and pure data for the hybrid wavelet modified GMDH model.

27.2.3 Hybrid Wavelet Modified GMDH Model

The main focus of this study is to construct a hybrid wavelet modified GMDH model for the monthly CPO price forecasting. The significant of the proposed hybrid model is to optimize in terms of computational resource and accuracy in comparison with individual GMDH and ARIMA model. In the proposed hybrid model, the complex data of the monthly crude palm oil price are effectively managed by the modification of conventional GMDH model and then hybridized with the discrete wavelet decomposition technique to enhance the forecasting accuracy. Radial basis, sigmoid, polynomial and tangent functions are selected as transfer functions and simultaneously integrated into GMDH for modified GMDH model. The advantages of these transfer functions are that the output can be easily determined for any given inputs and pick up the best and correct model to overcome the non-linearity of the data, which clarify the relationship between the given inputs and a proper output. It is also helpful to reduce the computational burden and consume less time on forecasting analysis, which makes it suitable for practical applications as compared to conventional GMDH model. The discrete wavelet transform (DWT) is selected as preprocessed clean and pure data, because using discrete wavelet transform the complex data of monthly palm oil price are decomposed into different sub series in such a way that error criteria is minimized. Once the data is trained after the sequential application of wavelet analysis, it is preferred to use it as an input to modified GMDH instead of conventional GMDH. As modified GMDH model is considered to be a better data trainer due to the advantages of simultaneous use of four transfer function for non-linear data. The result of the hybrid wavelet modified GMDH model for preprocessed clean and pure data is more precise than the network, which is trained for raw signal or noisy data.

The simplified structure of hybrid wavelet-modified GMDH model, in which the discrete wavelet analysis is linked as an input to the modified GMDH model is shown in Fig. 27.3. In the simplified structure of hybrid wavelet-modified GMDH model, first the number of input observations of complex data of monthly CPO price are determined. Then we have decomposed the complex data of monthly palm oil price using discrete wavelet transform with the help of MATLAB simulation tool. Next for the efficiency of the developed model, it is important to select the most effective wavelet components. The most effective wavelet components are determined by using the co-efficient of determination (R^2). The sum of the most effective components (detail (Ds) and approximation (As) coefficients) are constructed as a new input,

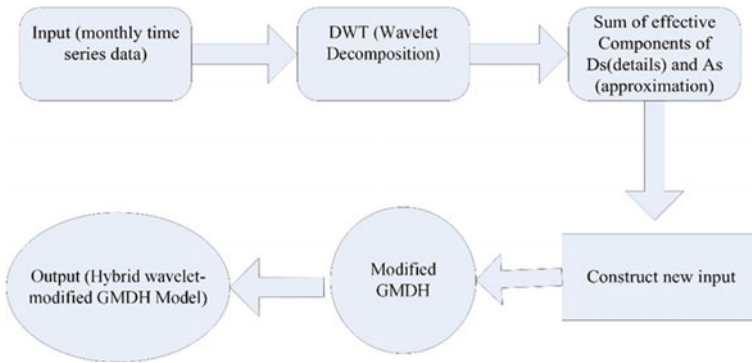


Fig. 27.3 The structure of hybrid wavelet-modified GMDH

which are used as an input to modified GMDH and thus the hybrid wavelet modified GMDH model has developed.

27.2.4 ARIMA Model

Autoregressive Integrated Moving Average (ARIMA) is a Box and Jenkins statistical model used for non-stationary time series, which is required to achieve stationary with the combination of AR and MA processes [38]. The AR components represent weighted sum of the past values, MA components for the weighted sum of pass errors and I stands for integrated [39]. The generalized form of full ARIMA (p, d, q) can be expressed as follow [40, 41].

$$x_t = \delta + \varphi_1 x_{t-1} + \varphi_2 x_{t-2} + \dots + \varphi_p x_{t-p} + a_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} \tag{27.2}$$

where x_t and a_t are the actual (current) value and random error at time period t respectively, while φ_i ($i = 1, 2, \dots, p$) AR parameters of the model and θ_j ($j = 0, 1, 2, \dots, q$) MA parameters of the model, p and q are integers and often referred to as orders of the model. The Box-Jenkins technique model has four important steps [42]: (i) Identification step (ACF and PACF) (ii) parameter estimation (iii) diagnostic checking to build ARIMA model (iv) forecasting step. For model identification, the appropriate order of the tentative models is identified through visual inspection of both ACF and PACF plot. The appropriate order is also supported by the lowest value of Akaike information criterion (AIC) for model identification step. The Augmented Dickey Fuller (ADF) test is used to test the stationarity of the monthly CPO price. The summary of the comparative study of hybrid GMDH neural network with individual and ARIMA model for the monthly CPO price of Malaysia is shown in Fig. 27.4.

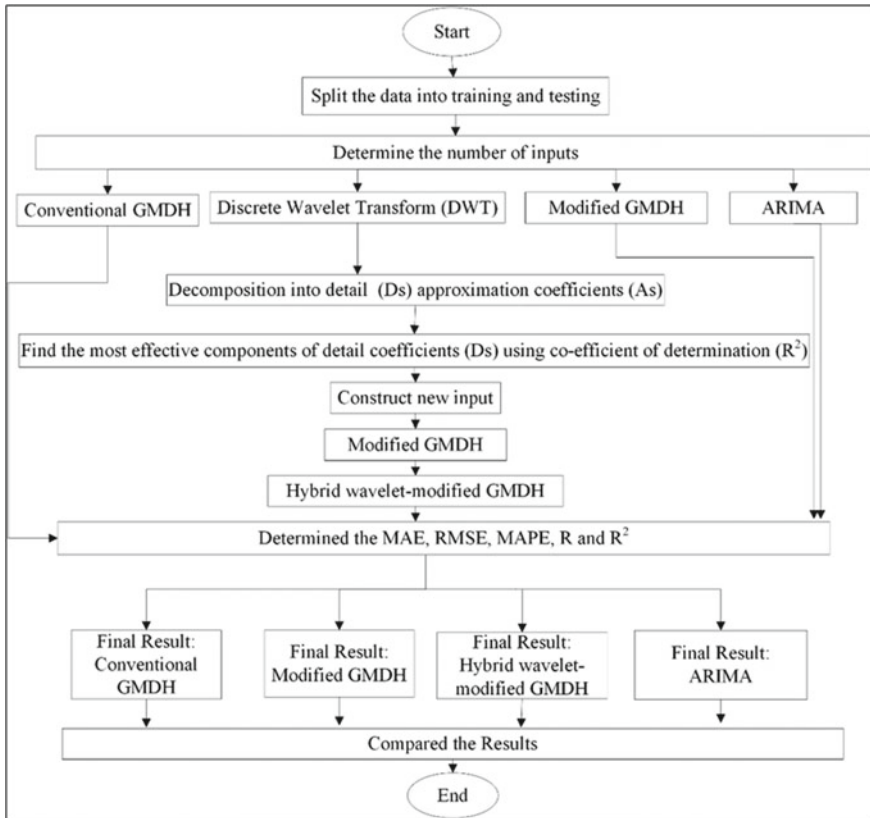


Fig. 27.4 Comparative study of Hybrid GMDH neural network with individual and ARIMA model

27.3 Performance Measurement

The following important statistical measures of errors are used to determine the performance of prediction of each model, which are given by the following formulas.

$$MAE = \frac{1}{N} \sum_{t=1}^N |y_t - \hat{y}_t| \tag{27.3}$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^N (y_t - \hat{y}_t)^2} \tag{27.4}$$

$$MAPE = \left(\frac{1}{N} \sum_{t=1}^N \left| \frac{y_t - \hat{y}_t}{y_t} \right| \right) \times 100\% \tag{27.5}$$

Correlation coefficient (R),

$$R = \frac{\frac{1}{N} \left[\sum_{t=1}^N (y_t - \mu_{y_t})(\hat{y}_t - \mu_{\hat{y}_t}) \right]}{\sqrt{\frac{1}{N} \sum_{t=1}^N (y_t - \mu_{y_t})^2} \sqrt{\frac{1}{N} \sum_{t=1}^N (\hat{y}_t - \mu_{\hat{y}_t})^2}} \tag{27.6}$$

Coefficient of determination (R^2).

$$R^2 = \left[\frac{\frac{1}{N} \left[\sum_{t=1}^N (y_t - \mu_{y_t})(\hat{y}_t - \mu_{\hat{y}_t}) \right]}{\sqrt{\frac{1}{N} \sum_{t=1}^N (y_t - \mu_{y_t})^2} \sqrt{\frac{1}{N} \sum_{t=1}^N (\hat{y}_t - \mu_{\hat{y}_t})^2}} \right]^2 \tag{7}$$

Here,

N = Number of data points.

y_t = Observed values at the time t .

μ_{y_t} = mean of the observed values.

\hat{y}_t = Forecasted values at the time t .

$\mu_{\hat{y}_t}$ = mean of forecasted value.

27.4 Results and Discussion

The results and discussion part shows the fitting of the monthly CPO price data of Malaysia into four different configurations. In the first part, conventional GMDH is applied on the monthly CPO price data. In second part, modified GMDH, in which the transfer functions are employed into GMDH model. In the third part for hybrid GMDH type neural network, the new input of monthly CPO price data obtained from the discrete wavelet analysis is linked to the modified GMDH model, which is known as hybrid wavelet modified GMDH model. In the fourth and last configuration a standard Box-Jenkins ARIMA (Autoregressive Integrated Moving Average) is employed on the monthly CPO price data for comparison. The sample data involved of 443 observations of monthly CPO price of Malaysia. The measurement data covered a duration of 37 years ranging from Jan, 1983 to Nov, 2019 and collected from the Malaysian Palm Oil Board (MPOB). The time series data of monthly crude palm oil price are measured in USD/Metric-Ton. The average monthly CPO price is 572.2, whereas maximum monthly price is 1292.0 and minimum price was 197.0.

Figure 27.5a shows the predicted (fitted) values after using conventional GMDH model for monthly crude palm oil price. The predicted values of conventional GMDH model follow the same pattern of the original monthly CPO price of time series data. The forecasted values of Conventional GMDH model for monthly CPO price of time series data are given in Table 27.2. Figure 27.5b presents the scatter plot of the optimal model between observed values (original/actual) and predicted values of conventional GMDH model. The scatter plot shows a linear relationship with good correlation coefficient (R) which is 0.9785 and coefficient of determination (R^2) which is 0.9570. The forecasting accuracy measurements of conventional GMDH for monthly CPO price are determined by using the important statistical measure of

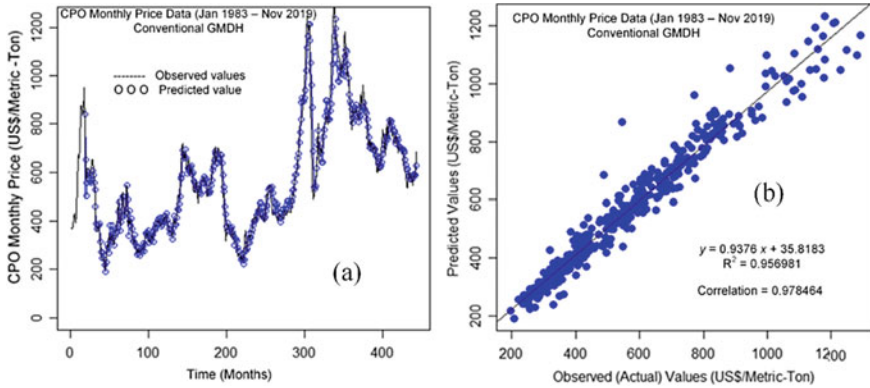


Fig. 27.5 Conventional GMDH **a** monthly CPO price versus time and **b** scatter plot

Table 27.2 Forecasted values of the predicted models for monthly CPO price of Malaysia

Forecast values of monthly CPO price (Jan 1983–Nov 2019)					
Point forecast	444	445	446	447	448
Months	Dec 2019	Jan 2020	Feb 2020	March 2020	April 2020
Conventional GMDH	670	677	689	771	708
Modified GMDH	705.0094	711.9125	714.5999	715.8732	716.3976
Hybrid wavelet modified GMDH	672.6053	676.9862	674.2843	669.0750	663.2505
ARIMA (7, 1, 7)	745.3897	765.4715	788.9943	808.6615	805.2089

errors such as mean absolute error (MAE), root mean square error (RMSE), MAPE (mean absolute percentage error) and performance measurements such as correlation coefficient (R) and coefficient of determination (R^2). The statistical measures of errors (accuracy measurements) of conventional GMDH model are measured carefully and given in Table 27.6.

Figure 27.6a shows the predicted (fitted) values after using modified GMDH model for monthly CPO price. In modified GMDH type neural network, transfer functions such as sigmoid, radial, tangent and polynomial are employed to improve the prediction accuracy of conventional GMDH using R programming language and GMDH package. The predicted values of modified GMDH model are in good understanding with the original monthly CPO price of time series data. The forecasted values of modified GMDH model for monthly CPO price of time series data are given in Table 27.2. Figure 27.6b presents the scatter plot of the optimal model between observed values (original/actual) and predicted values of modified GMDH model. A scatter plot shows a linear relationship with good correlation coefficient (R) which is 0.9840 and coefficient of determination (R^2) which is 0.9682. The forecasting accuracy measurements of modified GMDH for monthly CPO price are determined by using the important statistical measures of errors such as mean absolute error

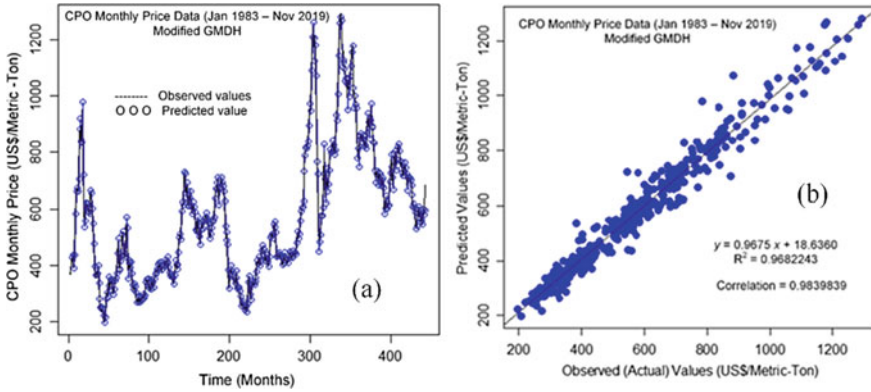


Fig. 27.6 Modified GMDH a monthly CPO price versus time and b scatter plot

(MAE), root mean square error (RMSE), MAPE (mean absolute percentage error) and performance measurements such as correlation coefficient (R) and coefficient of determination (R^2). The statistical measures of errors (accuracy measurements) of modified GMDH model are measured and given in Table 27.6.

In order to enhance the performance of the model, a hybridization of two different models are made by combining a discrete wavelet transform (DWT) and modified GMDH model. The wavelet analysis has been first applied to decompose the original monthly CPO price data after the simulation code of Mat-lab programming language into approximate (As) and details (Ds) components. As the decomposition of data is a finite step and to obtain acceptable results by using a certain level decomposition using the formula $M = \text{int}[\log(N)]$, where N represents length of monthly CPO price time series data, int represent for integer numbers and M represents level of decomposition [43]. The wavelet analysis used in this study based on 3 level decomposition. Figure 27.7a represents the plot of original monthly CPO price data series, where Fig. 27.7b–d have shown the details components (D1, D2 & D3) and Fig. 27.7e has shown the final approximate component (A3) as discussed in the methodology section. Figure 27.7f represents the summation plot of the most effective components (D2 + D3 + A3), which are determined by using correlation coefficient (R) or coefficient of determination (R^2) as shown in Table 27.3. Figure 27.7f has been constructed as a new model input after the summation of most effective components of wavelet analysis, which are used as a clean and pure data for the hybrid wavelet modified GMDH model.

Figure 27.8a shows the predicted (fitted) values after using hybrid wavelet modified GMDH model of monthly crude palm oil price. The predicted values of hybrid wavelet modified GMDH model are in good understanding with the original monthly CPO price of time series data. The forecasted values of hybrid wavelet modified GMDH model for monthly CPO price of time series data are given in Table 27.2. Figure 27.8b presents the scatter plot of the optimal model between observed values (original/actual) and predicted values of hybrid wavelet modified GMDH model. The

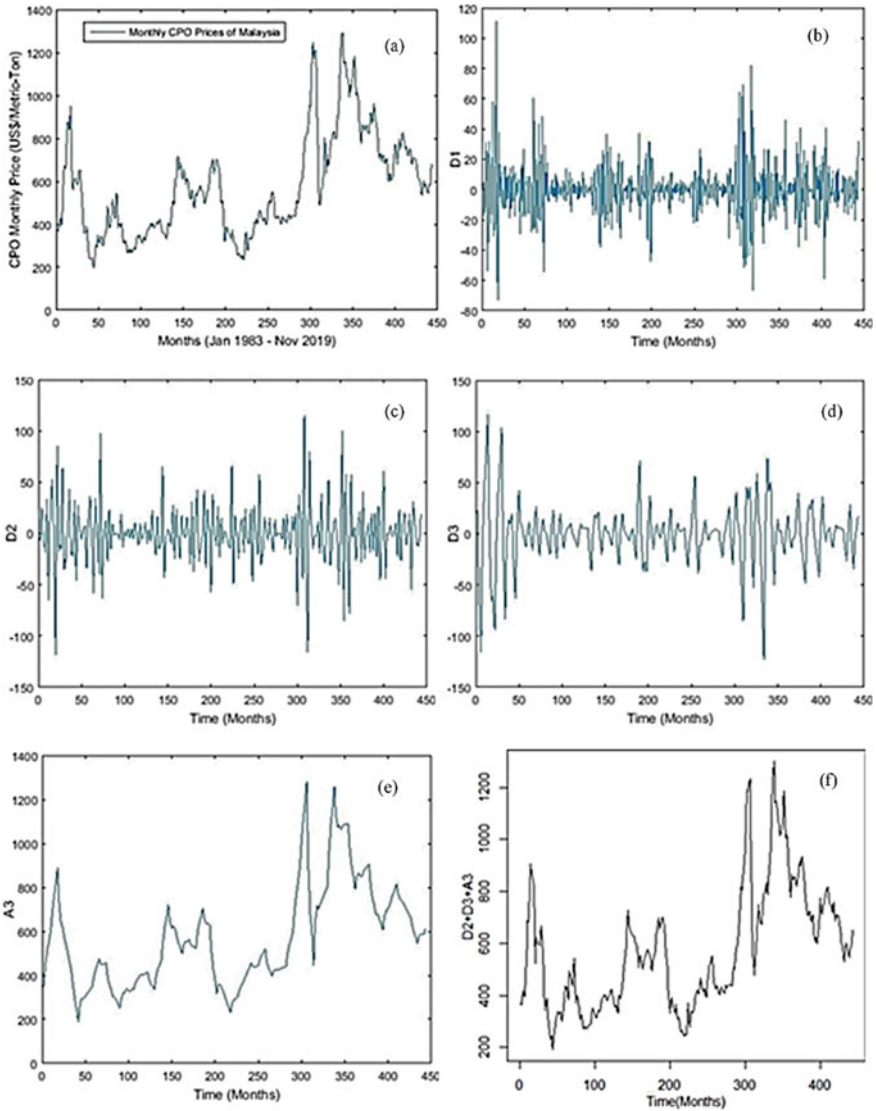


Fig. 27.7 Discrete wavelet decomposition of monthly CPO price into details (Ds) and approximate (As) components

Table 27.3 The correlation coefficient between discrete wavelet component and actual CPO price (Jan 1983–Nov 2019)

Wavelet components	R	R^2	$R^2(\%)$
D1	0.0851	0.0072	0.72
D2	0.1257	0.0158	1.6
D3	0.1350	0.0182	1.8
A3	0.9832	0.9667	96.67

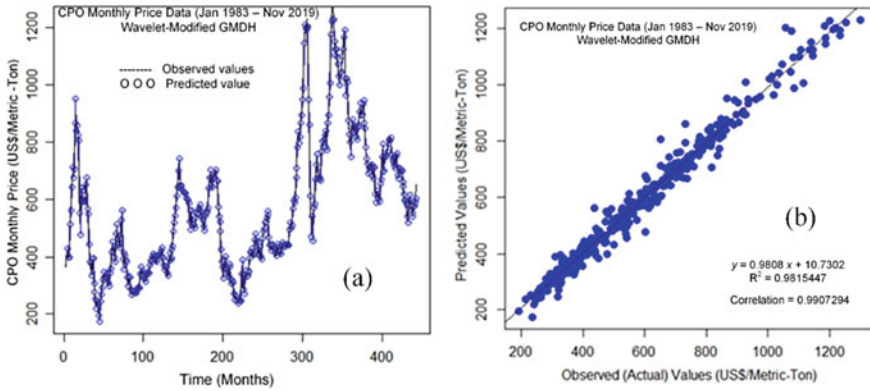


Fig. 27.8 **a** Crude palm oil prices versus time of hybrid wavelet-modified GMDH model, **b** scatter plot of hybrid wavelet-modified GMDH model

scatter plot shows a linear relationship with good correlation coefficient (R) which is 0.9907 and coefficient of determination (R^2) which is 0.9815. The forecasting accuracy measurements of hybrid wavelet modified GMDH model for monthly CPO price are determined by using the important statistical measures of errors such as mean absolute error (MAE), root mean square error (RMSE), MAPE (mean absolute percentage error) and performance measurements such as correlation coefficient (R) and coefficient of determination (R^2). The statistical measure of errors of hybrid wavelet modified GMDH model are measured and given in Table 27.6.

The comparative study of the hybrid wavelet modified GMDH with individual GMDH models and also with Box-Jenkins ARIMA model have been performed. The Box-Jenkins ARIMA (p, d, q) model of monthly CPO price data of Malaysia from January 1983 to Nov 2019 shows a non-stationary trend with time. To handle the non-stationary issue, the unit root test known as augmented Dickey-Fuller (ADF) test is used. ADF test is most frequently used to check the stationarity of time series data. From the ADF results of the monthly CPO price data given in Table 27.4, it has been clearly shown that the data set of time series obtained stationarity after taking the first difference, which represents as the alternative hypothesis is accepted, while the null

Table 27.4 Unit root ADF Test for monthly CPO price data (Jan 1983–Nov 2019)

Data: $X_t = CPO$ For $d = 0$		
Dickey-fuller = -3.1829	Lag order = 7	p -value = 0.0909
Rejecting alternative hypothesis [(H_a) : stationary], accepting null hypothesis (H_0): non-stationary		
Data (overall): $X_t - X_{t-1} = CPO$ For $d = 1$ (1st difference p -value)		
Dickey-fuller = -8.508,	Lag order = 7	p -value = 0.01
Accepting alternative hypothesis [(H_a) : stationary], rejecting null hypothesis [(H_0) : non-stationary]		

hypothesis is rejected. Thus, there is no need for further differencing the time series and we have adopted $d = 1$ for our ARIMA (p, d, q) model. ARIMA (7, 1, 7) models are build based on ACF and PACF plots with the lowest Akaike Information Criteria (AIC) value. Figure 27.9a shows the fitted (predicted) values of statistical ARIMA (7, 1, 7) model for monthly CPO price of time series data. The forecasted values of ARIMA (7, 1, 7) model for monthly CPO price of time series data are given in Table 27.2. Figure 27.9b presents the scatter plot of the optimal model between observed values (original/actual) and predicted values of ARIMA (7, 1, 7) model. The scatter plot shows a linear relationship with good correlation coefficient (R) which is 0.9859 and coefficient of determination (R^2) which is 0.9719. The estimated coefficient parameters and their standard errors of fitted ARIMA (7, 1, 7) model for the monthly CPO price are shown in Table 27.5. The table has been shown the selected values of good estimation coefficient and lowest AIC value for ARIMA (7, 1, 7) model with ($p = 7, d = 1$ and $q = 7$) and therefore the selected model could be most suitable predictive model for forecasting the future values of monthly CPO price time series data. The statistical measure of errors (accuracy measurements) of ARIMA (7, 1, 7) model are measured carefully and given in Table 27.6. The diagnostic test of ARIMA model for the monthly CPO price is shown in Fig. 27.10. The time series plot of the standardized residuals indicates that there is no trend in the residuals, while ACF of the residuals shows no significant autocorrelations which indicates a good result. Based on the Ljung-Box statistics, the p-value of ARIMA (7, 1, 7) model greater than 0.05 indicates that there is no autocorrelation in the residuals and the models are adequate.

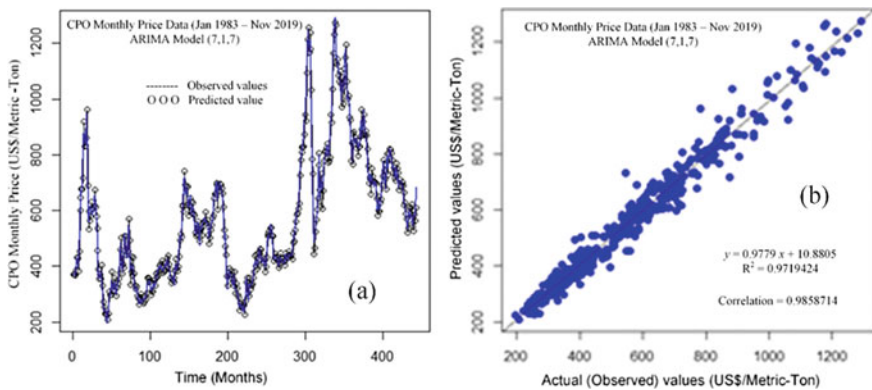


Fig. 27.9 a Crude palm oil prices versus time of ARIMA (7, 1, 7) model and b scatter plot of ARIMA (7, 1, 7) model

Table 27.5 Coefficients, std. errors and other parameters for ARIMA (7, 1, 7) for monthly CPO price forecasting

ARIMA (7, 1, 7)									
Coefficients	AR1	AR2	AR3	AR4	AR5	AR6	AR7		
AR TERMS	0.4392	0.5600	-0.3148	- 0.4032	0.4452	-0.4155	0.3961		
Std. error	0.1593	0.1757	0.1473	0.1704	0.1593	0.1751	0.1072		
P-value	0.0058325	0.0014418	0.021774	0.0179565	0.0065520	0.0193819	0.0002212		
Coefficients	MA1	MA2	MA3	MA4	MA5	MA6	MA7		
MA TERMS	-0.3139	-0.7614	-0.4078	0.7249	-0.4602	-0.2999	-0.4950		
Std. error	0.1597	0.1506	0.2021	0.1527	0.1831	0.1379	0.1299		
P-value	0.0417560	4.248e ⁻⁰⁷	0.04881	2.048e ⁻⁰⁶	0.0003779	0.0417560	0.0001381		

σ^2 estimated as 1527; log likelihood = -2248.88, AIC = 4527.76

Ljung-Box (p-value) = 0.9694
 ARIMA (7, 1, 7), AIC = 4527.76
 ARIMA (7, 1, 6), AIC = 4531.01
 ARIMA (7, 1, 8), AIC = 4529.03

Table 27.6 Comparison of forecasting accuracy (fit) of monthly CPO price of Malaysia

Accuracy (fit) of monthly CPO price data (Jan1983–Nov 2019)					
Model	MAE	RMSE	MAPE	<i>R</i>	<i>R</i> ²
Conventional GMDH	33.1849	48.4489	5.9850	0.9785	0.9570
Modified GMDH	28.8580	41.5307	5.2733	0.9840	0.9682
Hybrid wavelet modified GMDH	21.0506	31.5526	3.9860	0.9907	0.9816
ARIMA	28.2005	39.0313	5.1434	0.9859	0.9719

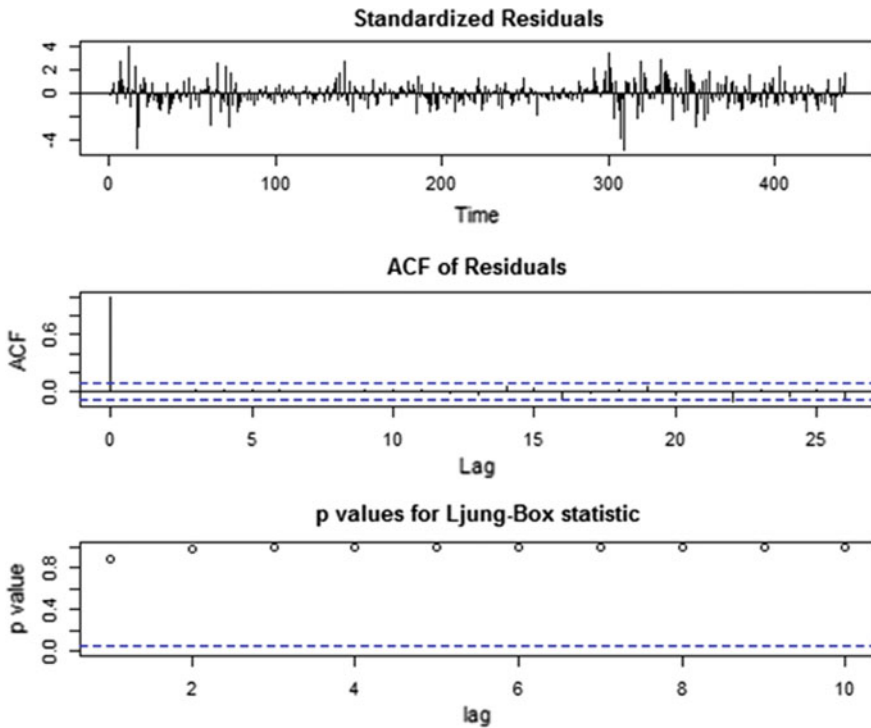


Fig. 27.10 Diagnostic test of ARIMA (7, 1, 7) model for monthly CPO price data (Jan 1983–Nov 2019)

27.5 Comparison of Forecasting Accuracy Measurements

The purpose of this study is to expand modeling and forecasting of monthly CPO price of Malaysia based on the forecasting accuracy measurements of hybrid GMDH neural network, individual GMDH and ARIMA models. The forecasted values and their accuracy measurements of monthly CPO price of Malaysia for hybrid GMDH neural network, individual GMDH and ARIMA model are given in Tables 27.2 and 27.6 respectively. Table 27.6 presents the comparison of forecasting accuracy of the

fitted models. The capabilities of hybrid GMDH neural network, individual GMDH and ARIMA models in modelling and forecasting the monthly CPO price are determined by using Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), MAPE, correlation coefficient (R) and coefficient of determination (R^2). The results of hybrid wavelet modified GMDH models have shown an improvement over statistical ARIMA model with 25.35%, 19.16% and 22.50% reduction in MAE, RMSE and MAPE values respectively. The obtained results of the comparison show that the hybrid wavelet-modified GMDH neural network is more accurate and performed better than ARIMA model and individual GMDH neural network.

27.6 Conclusion

Modelling and forecasting of monthly CPO price of Malaysia based on the forecasting accuracy measurements of hybrid GMDH neural network, individual GMDH and ARIMA models have been investigated. The discrete wavelet transform has been successfully selected as a pre-processed clean and pure data of model inputs and was linked with modified GMDH neural network. Modified GMDH model has been performed a better data trainer due to the advantages of simultaneous use of transfer functions for non-linear data. The accuracy measurements of hybrid GMDH, individual GMDH and ARIMA models have been determined by using MAE, RMSE, MAPE, R and R^2 . The results of hybrid GMDH type neural network have shown an improvement over statistical ARIMA model with 25.35%, 19.16% and 22.50% reduction in MAE, RMSE and MAPE values respectively. The obtained results of the comparison show that the hybrid wavelet-modified GMDH neural network has performed better result than ARIMA model and individual GMDH neural network. The proposed hybrid wavelet modified GMDH model has been proved more effective and efficient way to improve the forecasting accuracy and will be suitable for practical applications.

Acknowledgements The authors would like to express their appreciation for the support of the sponsors with Project Vote No: U305 by Research and Innovation Centre (R&D), Research, Innovation, Commercialization and Consultancy Office (ORICC) University Tun Hussein Onn Malaysia. The author would also like to thanks, Centre of Graduate Studies (CGS), Universiti Tun Hussein Onn Malaysia for the financial support in making the project a success. Both supports is gratefully acknowledged.

References

1. D Chalil R Barus 2018 The sustainability of Indonesia Pakistan palm oil supply chain J. Manaj. Dan Agribisnis. 15 136 142 <https://doi.org/10.17358/jma.15.2.136>
2. MK Lam KT Tan KT Lee AR Mohamed 2009 Malaysian palm oil: surviving the food versus fuel dispute for a sustainable future Renew. Sustain. Energy Rev. 13 1456 1464 <https://doi.org/10.1016/j.rser.2008.09.009>
3. MH Ahmad PY Ping N Mahamed 2014 Volatility modelling and forecasting of Malaysian crude palm oil prices Appl. Math. Sci. 8 6159 6169 <https://doi.org/10.12988/ams.2014.48650>
4. A.L. Mohamad Hanapi, M. Othman, R. Sökkalingam, H. Sakidin, Developed a hybrid sliding window and GARCH model for forecasting of crude palm oil prices in Malaysia. J. Phys. Conf. Ser. **1123** (2018). <https://doi.org/10.1088/1742-6596/1123/1/012029>
5. A. Bin Khamis, R. Hameed, M.E. Nor, N. Che Him, R. Mohd Salleh, S.N.A. Mohd Razali, Comparative Study on Forecasting Crude Palm Oil Price using Time Series Models, Sci. Res. J. VI (2018). doi:<https://doi.org/10.31364/scirj/v6.i12.2018.p1218588>
6. N Khalid HNA Hamidi S Thinagar NF Marwan 2018 Crude palm oil price forecasting in Malaysia: an econometric approach J. Ekon. Malaysia. 52 263 278
7. N.A. Ismail, B.A. Talib, A. Mokhtar, Export analysis of major commodities in Malaysia, IOP Conf. Ser. Earth Environ. Sci. **327** (2019). <https://doi.org/10.1088/1755-1315/327/1/012002>
8. F.F. Abu Hassan Asari, N.H. Abdul Rahman, E.A. Abdul Razak, B.A. Ahmad, Shabir, N.F.A. Harun, K. Jusoff, A time series analysis of the relationship between total area planted, palm oil price and production of Malaysian palm oil. World Appl. Sci. J. **12**, 34–40 (2011)
9. N.M. Abas, N. Junaini, A. Yaacob, S. Nurul, A. Mohd, N.M. Yusof, R.A. Azidin, An overview of primary sector in Malaysia, Int. J. Econ. Commer. Manag. United Kingdom Percent. **III**, 1–13 (2015)
10. A Khamis Z Ismail K Haron AT Mohammed 2005 Nonlinear growth models for modeling oil palm yield growth J. Math. Stat. 1 225 232 <https://doi.org/10.3844/jmssp.2005.225.232>
11. H Basheer A Khamis 2017 Forecasting of crude palm oil price using hybridizing wavelet and group method of data handling model, Malaysian J. Fundam. Appl. Sci. 13 642 648
12. Azme Khamis, Application of statistical and neural network model for oil palm yield study, Ph.D. Thesis Faculty Science Universiti Teknologi Malaysia (2005)
13. A. Khamis, N.A. Abdul Razak, M.A.A. Abdullah, A robust vector autoregressive model for forecasting economic growth in Malaysia, Malaysian. J. Fundam. Appl. Sci. **14**, 382–385 (2018). <https://doi.org/10.11113/mjfas.v14n3.1021>
14. A. Bin Khamis, N. Khalidah, K. Binti, Comparative Study On Estimate House Price Using Statistical And Neural Network Model, Int. J. Sci. Technol. Res. 3 (2014) 126–131
15. S Sufahani N Che-Him A Khamis MS Rusiman NA Arbin CK Yee IN Ramli NA Suhaimi SS Jing ZA Azmi 2017 Descriptive statistics with box-jenkins and marketing research for jewellery company in Malaysia Far East J. Math. Sci. 101 2151 2161 <https://doi.org/10.17654/MS101102151>
16. MEB Nor MS Rusiman SF Sufahani MAA Abdullah SA Bataraja S Saharan 2018 Deseasonalisation in electricity load forecasting Int. J. Eng. Technol. 7 448 <https://doi.org/10.14419/ijet.v7i4.30.22364>
17. S Publications 2012 Parameters estimate of autoregressive moving average and autoregressive integrated moving average models and compare their ability for inflow forecasting Mohammad Valipour, Mohammad Ebrahim Banihabib and Seyyed Mahmood Reza Behbahani College of Abureyhan J. Math. Stat. 8 330 338 <https://doi.org/10.3844/jmssp.2012.330.338>
18. M Aamir A Shabir 2016 Modelling and forecasting monthly crude oil price of Pakistan: a comparative study of ARIMA, GARCH and ARIMA Kalman model AIP Conf. Proc. 1750 2365 2371 <https://doi.org/10.1063/1.4954620>
19. T. Kondo, a. S. Pandya, J.M. Zurada, Logistic GMDH-type neural networks and their application to the identification of the X-ray film characteristic curve, in IEEE SMC'99 Conf. Proceedings. 1999 IEEE Int. Conf. Syst. Man, Cybern. (Cat. No.99CH37028). 1 (1999) 437–442. <https://doi.org/10.1109/ICSMC.1999.814131>

20. A Darvizeh N Nariman-Zadeh H Gharababaei 2003 GMDH-type neural network modelling of explosive cutting process of plates using singular value decomposition Syst. Anal. Model. Simul. 43 1383 1397 <https://doi.org/10.1080/02329290290024358>
21. LI Qiumin T Yixiang Z Gaoxun 2014 An Elman model based on GMDH algorithm for exchange rate forecasting Manag. Sci. Eng. 8 1 6 <https://doi.org/10.3968/5666>
22. M. Najafzadeh, G.A. Barani, M.R. Hessami Kermani, GMDH based back propagation algorithm to predict abutment scour in cohesive soils. Ocean Eng. **59**, 100–106 (2013). <https://doi.org/10.1016/j.oceaneng.2012.12.006>
23. H Basheer A Khamis 2016 A hybrid group method of data handling (GMDH) with the wavelet decomposition for time series forecasting: a review ARPN J. Eng. Appl. Sci. 11 10792 10800
24. P.C. Parks, a. G. Ivakhnenko, L.M. Boichuk, B.K. Svetalsky, A self-organizing model of the British economy for control with optimal prediction using the balance-of-variables criterion. Int. J. Comput. Inf. Sci. **4**, 349–379 (1975). <https://doi.org/10.1007/BF00979374>
25. S. Ikeda, M. Ochiai, Y. Sawaragi, Sequential GMDH algorithm and its application to river flow prediction. IEEE Trans. Syst. Man, Cybern. **6**, 473–479 (1976)
26. M. Robinson, C., Mort, Predicting foreign exchange rates using neural and genetic models. Proc. 2nd Asian Control Conf. , Seoul, Korea **3**, 115–118 (1997). <https://doi.org/10.1017/CBO9781107415324.004>
27. T.H. Hayashi Isao, The fuzzy GMDH algorithm by possibility models and its application. Fuzzy Sets Syst. **36**, 245–258 (1990)
28. T. Kondo, A.S. Pandya, Identification of the multi-layered neural networks by revised GMDH-type neural network algorithm with PSS criterion, M.Gh. Negoita AI. KES 2004, LNAI 3214, © Springer-Verlag Berlin Heidelb. 1051–1059 (2004)
29. GC Onwubolu 2008 Design of hybrid differential evolution and group method of data handling networks for modeling and prediction Inf. Sci. (Ny) 178 3616 3634 <https://doi.org/10.1016/j.ins.2008.05.013>
30. A. Darvizeh, Hybrid genetic design of GMDH-type neural networks using singular value decomposition for modelling and prediction of the explosive cutting, **217** (n.d.) 779–790
31. S.J. Farlow, *Self-Organizing Methods in Modeling: {GMDH} Type Algorithms* (1984)
32. P Buryan GC Onwubolu 2011 Design of enhanced MIA-GMDH learning networks Int. J. Syst. Sci. 42 673 693 <https://doi.org/10.1080/00207720903225526>
33. L. Anastasakis, L. Anastasakis, N. Mort, N. Mort, The development of self-organization techniques in modelling: a review of the group method of data handling (gmdh), Dep. Autom. Control Syst. Eng. Univ. Sheff. Mappin St, Sheffield, S1 3JD, United Kingdom. (2001) 1–38
34. D Kim SJ Seo GT Park 2009 Hybrid GMDH-type modeling for nonlinear systems: synergism to intelligent identification Adv. Eng. Softw. 40 1087 1094 <https://doi.org/10.1016/j.advengsoft.2009.01.029>
35. V. Lebedev, *ITHEA Time Series Prognosis of GDP with the System GMDH-Shell (experimental work) Artificial Intelligence Driven Solutions to Business and Engineering Problems*, ITHEA 32–36 (2008)
36. O Dag C Yozgatligil 2016 GMDH: an R package for short term forecasting via GMDH-type neural network algorithms R J. 8 379 386
37. T Kondo J Ueno S Takao 2013 Hybrid multi-layered GMDH-type neural network using principal component regression analysis and its application to medical image diagnosis of liver cancer Procedia Comput. Sci. 22 172 181 <https://doi.org/10.1016/j.procs.2013.09.093>
38. R. Dalinina, Introduction to forecasting with ARIMA in R introduction to time series forecasting. Data Sci. 1–24 (2018). <https://www.datascience.com/blog/introduction-to-forecasting-with-arima-in-r-learn-data-science-tutorials>
39. S Ahmad HA Latif 2011 Forecasting on the crude palm oil and kernel palm production: seasonal ARIMA approach, 2011 IEEE Colloq Humanit. Sci. Eng. CHUSER 2011 2 939 944 <https://doi.org/10.1109/CHUSER.2011.6163876>
40. M.S. Voss, X. Feng, {ARMA} Model selection using particle swarm optimization and {AIC} criteria, IFAC 15th Trienn. World Congr. **2** (2002)

41. OY Martins 2011 ARMA modelling of Benue river flow dynamics: comparative study of PAR model Open J. Mod. Hydrol. 01 1 9 <https://doi.org/10.4236/ojmh.2011.11001>
42. M Yousaf S Zahir M Riaz S Muhammad 2020 Statistical analysis of forecasting COVID-19 for upcoming month in Pakistan Chaos, Solitons Fractals 138 1 4
43. V Nourani MT Alami MH Aminfar 2009 A combined neural-wavelet model for prediction of Ligvanchai watershed precipitation Eng. Appl. Artif. Intell. 22 466 472 <https://doi.org/10.1016/j.engappai.2008.09.003>

Chapter 28

Analyzing of Traffic Accidents in Skudai City Using Stochastic Models



Teng Mun Jing and Rohayu Mohd Salleh 

Abstract This study aims to investigate the pattern of traffic accidents in Skudai, Johor. Due to the lack of comprehensive public transportation development in Skudai area, the number of registered vehicles have increased dramatically, also the number of traffic accidents. The data is provided via Ibu Pejabat Polis Daerah Johor Bahru (Utara), Skudai from 2015 to 2019. Data scopes were divided into number of daily accidents, number of daily fatal accidents, and number of daily fatal victims. Data visualization and descriptive statistics were used to analyse the pattern and summarization of the daily accidents. Thus, estimating accident parameters such as daily accident rate, daily accident hazard level, daily fatal accident probability, and daily fatal victim rates are critical for traffic management. Three stochastic models are used to estimate those parameters: the Generalized Poisson (GP) distribution, the Generalized Poisson-Quasi Binomial (GPQB) distribution, and the Generalized Poisson-Generalized Poisson (GPGP) distribution. The goodness-of-fit test was used to evaluate if the models were statistically significant fit to the data. The findings demonstrate that while the general accident rate and accident hazard level increased gradually, the likelihood of fatal accidents and the rate of fatal victims dropped in Skudai. The GP model was only significantly fitted in 2017 and 2019, whereas the GPGP model was only significant fitted in 2017. The GPQB model resulted that there was no significantly fitted in all of the years.

T. M. Jing

Department of Mathematics and Statistics, Faculty of Applied Sciences and Technology,
Universiti Tun Hussein Onn Malaysia, 84600 Pagoh, Johor, Malaysia

e-mail: munjing1015@gmail.com

R. M. Salleh (✉)

Data Analytics, Sciences and Modeling Focus Group, Universiti Tun Hussein Onn Malaysia,
84600 Pagoh, Johor, Malaysia

e-mail: msrohayua@uthm.edu.my

28.1 Introduction

28.1.1 *Background of the Study*

Nowadays, road transportation plays an essential role in economic development, logistics, tourism industry, and others [1]. The advantages of this mode of transportation such as lower investment required and cheaper cost, but the negative impact is a higher number of road accidents which may lead the problems in socio-economic and public health. Road accidents may affect by the larger population and higher registered vehicles [2]. In Johor, there has a higher number of road accidents from 2009 to 2018 [3].

Skudai is part of the developing area in Johor and a suburb in Iskandar Puteri, Johor Bahru. The lack of railway transportation development such as LRT, MRT, or Monorail in Skudai forced people to own their vehicles in order to arrive at their destination. This scenario will cause an increasing number of registered vehicles, traffic jams, or road accidents in this society. These factors would make the traffic situation more complicated to determine the pattern of daily road accidents.

It is important to understand the pattern of traffic accident events and subsequently decrease the number of road accidents. Therefore, this study aims to have a better understanding of the pattern of unexpected phenomena for daily traffic accidents. In order to have a more comprehensive understanding of the occurrence of daily accidents, this study is focused on the accident parameters, which are daily accident rate, daily accident hazard level, probability of daily fatal accidents, and daily fatal victim rate by using the method of moments to estimate the parameters. Through this study, there is vital for the authority in traffic management to be able to take necessary actions for reducing the quantity and quality of traffic accidents based on the condition of Skudai [4].

28.2 The Objectives of the Study

In this study, the stochastic models would be applied to determine the behaviour of the parameters. Using the stochastic models, at least three objectives are expected to achieve. The first is to explain accurately the accidents parameters, to estimate those parameters for Skudai city based on actual data, and to measure the accuracy of the model by conducting a goodness of fit test. The remaining of this study is organized as follows. Section 28.2 presents the research methodology used to perform the analysis along with the dataset. Section 28.3 presents the results and Sect. 28.4 concludes the study with some direction for future work.

Table 28.1 Description of variables

Group	Variables	Descriptions
X	Year	Year for the daily accidents occurred
	Number of accidents, X	Number of daily accidents occurred in one day per year
	Number of days	Number of days that occurred the accidents in the following year
Y	Year	Year for the daily fatal accidents occurred
	Number of fatal accidents, Y	Number of daily fatal accidents occurred in the daily accidents
	Number of days with X	Number of days that occurred the fatal accidents in the daily accident
Z	Year	Year for the daily fatal victims occurred
	Number of fatal victims, Z	Number of daily fatal victims occurred in the daily accidents
	Number of days with X	Number of days that occurred fatal victims in the daily accident

28.3 Methodology

28.3.1 Data Collection

In this study, Skudai daily accidents data was collected via Ibu Pejabat Polis Daerah Johor Bahru (Utara), Skudai from 2015 to 2019. The data categorized into three groups X, Y and Z. Data in group X is the number of daily accidents, data in group Y is the number of daily fatal accidents, and data in group Z is the number of daily fatal victims. Table 28.1 shows the description of variables among the three tables.

28.4 Data Visualization

In this study, the data visualization technique is applied to produce the graphical chart for data in Table X, Table Y, and Table Z. The aim is to visualize the trends, outliers, and patterns of data from 2015 to 2019. The R software is used to present the results.

28.4.1 Descriptive Statistics

Descriptive statistics are the procedure to summarize the collection of data describing what occurred in the observed data [5]. In this study, descriptive statistics are applied

to produce the graphical charts and tables of summary. Two types of statistical measures applied in this study are the measures of central tendency and the measures of dispersion [6].

A measure of central tendency is to describe the data by determining the central position within a set of data. It may also be classed as summary statistics. The commonly used for the measure of central tendency are mean, mode, and median. For the measures of dispersion, there is a statistic that implies the variability, scatter or spread of the data. In this study, the standard deviation, kurtosis, and skewness are focused on explaining the pattern of occurrence of road accidents.

28.4.2 *Stochastic Modelling*

Stochastic modelling is a mathematical model that uses in the forecast, planning, control, or “understanding” the behaviour of the occurrence event to produce a specified outcome [7]. It also a quantitative description that reflected the natural phenomena [8]. Stochastic modelling due to the behaviour of traffic accidents that follows the stochastic process.

Three stochastic models adopted in this study, Generalized Poisson (GP) distribution, Generalized Poisson-Quasi Binomial (GPQB) distribution, and Generalized Poisson-Generalized Poisson (GPGP) distribution. Each of the distributions having different roles in explains the parameter. The GP distribution used for the daily accident rate and accident hazard level. Meanwhile, the GPQB distribution and the GPGP distribution for the probability of fatal accidents and daily fatal victim rate. Then, these three stochastic models tested in the Goodness-of-fit test to calculate the expected values for the dataset.

28.4.3 *Generalized Poisson Model*

This study is applied to a different Poisson model called Generalized Poisson Distribution (GPD) defined by [9]. In GPD, it involved a two-parameter family that will provide a better fit to data. This study is focused on the accident rate and accident hazards level as the parameters of the distribution. Let assume X is a discrete variable and denote the number of daily accidents during the interval period. If the study is focused on the accident rate parameter only, which is equivalent to the assumption that the occurrence of accidents in a random pattern, then it is an appropriate way to apply the Poisson process in this study [4]. Let assume that all the possible outcome of X with the space $A = \{0, 1, 2 \dots\}$. If the two parameters as the accidents rate (θ) and accident hazard level (λ), X has followed the probability density function. Thus, the notation can be written as $X \sim GP(\theta, \lambda)$. Hence, the function as follows:

$$f(x) = \frac{\theta(\theta + x\lambda)^{x-1} e^{-(\theta+x\lambda)}}{x!} \tag{28.1}$$

where x is in A and $\theta > 0$ [9]. The mean and variance of X are respectively,

$$\mu_x = \theta(1 - \lambda)^{-1} \text{ and } \sigma_x^2 = \theta(1 - \lambda)^{-3} \tag{28.2}$$

There have some relationships of the hazard level of accident, λ with the different situations. When $\lambda = 0$, then it implies that the phenomenon as “no effort” to be done as avoiding an accident. Meanwhile, the occurrence of the accident is in a random pattern. When $\lambda < 0$, then it implies that there has an effort from the road users to reduce an accident occurred. Meanwhile, there has a certain pattern of road users’ behaviour to avoid the occurrence of accidents. When $\lambda > 0$, it implies that there is no effort to reduce an accident from the road users. In other words, if the value of $\lambda > 0$ is larger, then the level of many hazards in increasing the number of traffic accidents is higher.

28.4.4 Probability of Fatal Accidents

Probability is a chance of the occurrence of events [10]. This study is concerned with the probability of fatal accidents. This is because the fatal accident is one of the leading deaths in Malaysia. For the previous illustration of X , let say that the random variable X is $X = x$, where x is the number of days with X (number of accidents), and suppose let Y denote as the number of fatal accidents. The interval of Y is $B = \{0, 1, 2, \dots, x\}$. It indicated that the number of fatal accidents is based on daily accidents that occurred. Hence, this section will discuss the probability of fatal accidents with the distribution of the number of fatal accidents, Y among the number of daily accidents, X . Let Y_i defined as an indicator of random variables associated with the i th accident as follows:

$$Y_i = \begin{cases} Y_i = 1, & \text{if the } i\text{th accident is fatal} \\ Y_i = 0, & \text{if the } i\text{th accident is non-fatal} \end{cases}$$

where if $P(Y_i = 1) = p$ and $P(Y_i = 0) = q = 1 - p$. The joint probability density function $\Phi(x, y)$ by [11] as follows:

$$\phi(x, y) = \theta^2 p q (\theta p + y\lambda)^{y-1} (\theta q + (x - y)\lambda)^{x-y-1} \frac{\exp(-(\theta + x\lambda))}{(x - y)! y!} \tag{28.3}$$

where x in A , y in B , $0 < p < 1$, $\theta > 0$ and $\theta + x\lambda > 0$. Equation 28.3 consists of three parameters θ , λ , and p and this bivariate distribution of X and Y is called Generalized Poisson-Quasi Binomial (GPQB) distribution. Thus, this can briefly be written as GPQB (θ, λ, p) . The mean and variance of Y are respectively,

$$\mu_Y = \theta p(1 - \lambda)^{-1} \text{ and } \sigma_Y^2 = \theta p(1 - \lambda)^{-3} \tag{28.4}$$

28.4.5 Fatal Victim Rate

The fatal victim is from fatal accidents. This study concerned the daily accidents of $X = x$ in a given interval of time which have been mentioned in the previous subsection of this study. Let consider Z_i denotes as the random variables associated with the rate of fatal victims in the i th accident where $i = 0, 1, 2, \dots, x$ which was mentioned by [12]. In this subsection, Z_i of accident hazard level, λ will be assumed to follow the Poisson model. Thus, the interval of Z_i will be $A = \{0, 1, 2, \dots\}$ which is the same as the interval of X . Furthermore, let η represents the rate of fatal victims. So, Z_i can be briefly as $Z_i \sim GP(\eta, \lambda)$. The formula is referred to by [12] as follows:

$$\delta(x, z) = \frac{\theta \eta (\theta + x\lambda)^{x-1} (x\eta + z\lambda)^{z-1}}{(x - 1)!z!} \tag{28.5}$$

where x and z in A and $\theta + x\lambda + x\eta + z\lambda > 0$. Therefore, this bivariate distribution X and Z called Generalized Poisson-Generalized Poisson (GPGP) distribution. Since the joint distribution has parameters $\theta, \lambda,$ and $\eta,$ hence, the notation can be briefly as GPGP (θ, λ, η) . The mean and variance of Z are respectively,

$$\mu_z = \theta \eta (1 - \lambda)^{-2} \text{ and } \sigma_z^2 = \theta \eta (1 - \lambda + \eta) (1 - \lambda)^{-2} \tag{28.6}$$

28.4.6 Parameter Estimation

Parameter estimation is the process of making inferences about the parameters and estimating the parameters of a selected distribution. In this study, the parameter estimation would be used to estimate those parameters of daily accident rate (θ), accident hazard level (λ), daily fatal victim rate (η), and the probability of daily fatal accident (p). All the mean and variance of Eqs. 28.2, 28.4, and 28.6 are non-linear functions of those parameters. Since the maximum likelihood method of equations is more complicated to estimate those parameters, thus, the method of moment will be used instead of the maximum likelihood estimators. Let $X_1, X_2, \dots, X_m, Y_1, Y_2, \dots, Y_n$ and Z_1, Z_2, \dots, Z_k be random samples respectively from $X, Y,$ and Z of the size $m, n,$ and $k.$ Hence, the sample means and sample variance are as follows

$$\bar{X} = \frac{1}{m} \sum_{i=1}^m X_i, \bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i, \bar{Z} = \frac{1}{k} \sum_{i=1}^k Z_i \text{ and } S_X^2 = \frac{1}{m-1} \sum_{i=1}^m (X_i - \bar{X})^2 \tag{28.7}$$

The sample means are unbiased estimators of μ_X, μ_Y and μ_Z while the sample variances are also unbiased estimators of σ_Y^2 and σ_Z^2 . In order to apply the method of moment, the statistics need to construct with the following equations first as follows:

$$\bar{X} = \frac{\hat{\theta}}{1 - \hat{\lambda}}, S_X^2 = \frac{\hat{\theta}}{(1 - \hat{\lambda})^3}, \bar{Y} = \frac{\hat{\theta} \hat{p}}{1 - \hat{\lambda}}, \bar{Z} = \frac{\hat{\theta} \hat{\eta}}{(1 - \hat{\lambda})^2} \tag{28.8}$$

Therefore, Eq. 28.8 is obtained from the estimators. The moment estimators are referred to by [12] as follows:

$$\hat{\theta} = \bar{X}(1 - \hat{\lambda}), \hat{\lambda} = 1 - \frac{\sqrt{\bar{X}}}{s_x}, \hat{\eta} = \frac{\bar{Z}(1 - \hat{\lambda})^2}{\hat{\theta}}, \text{ and } \hat{p} = \frac{\bar{Y}}{\bar{X}} \tag{28.9}$$

28.4.7 Goodness of Fit Test

The goodness-of-fit test is a non-parametric test. The purpose of this test is to figure out whether the model is significantly fits the observed data. This study used the Goodness-of-Fit test for the distribution of $X, Y,$ and $Z,$ which are the basic models, by using the chi-square distribution to carry out the results. The goodness-of-fit test is examined the GP model which is the distribution of daily accidents, X as the basic model. The GPQB model is tested which is the distribution of fatal accidents, Y as the basic model. For the GPGP model, the goodness of fit test is conducted for the distribution of Z which is the fatal victims as the basic model. The values of parameters are substituted into the basic model of the equations to obtain the estimate of the probability values in each year of observed data. The expected values have used the estimate of probability values multiplied by the number of days in each year. The test statistics value was calculated by using the observed values and expected values. The decision will use the p -value to decide whether the null hypothesis is rejected. If the p -value is greater than 0.05, so it can be concluded that the null hypothesis cannot be rejected, the model is significantly fitted to the observed data.

28.5 Results and Discussions

28.5.1 Occurrence of Accidents

For the number of daily accidents, Fig. 28.1 shows the results of the bar chart to visualize the pattern of the accidents occurred. To support this finding in the bar chart, the descriptive statistics of the number of daily accidents is presented in Table 28.2.

Based on Fig. 28.1, there have frequently occurred road accidents above 10 days between 7 and 20 cases per day in 2015, 2016, and 2017. In 2018 and 2019, there have frequently occurred road accidents between 9 and 26 cases per day, which has

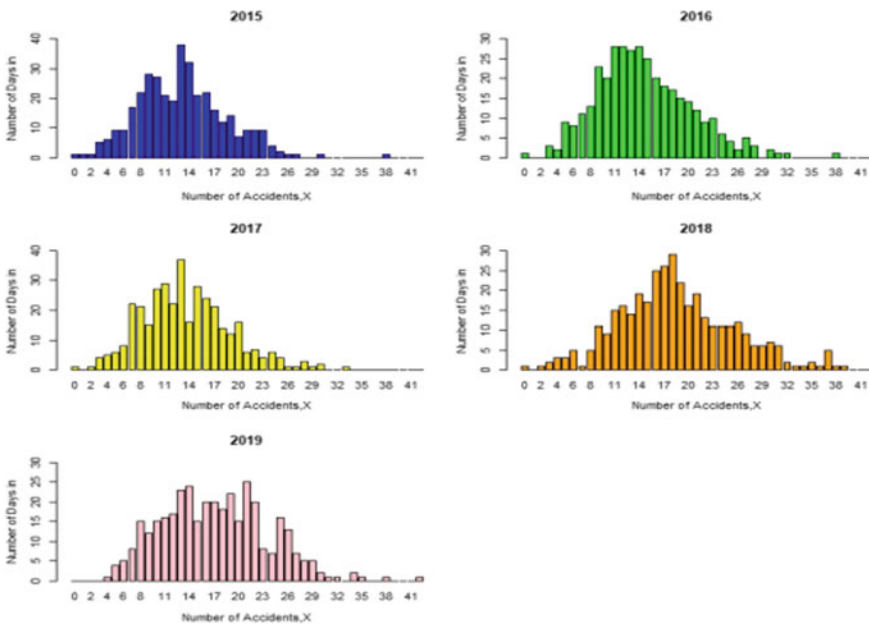


Fig. 28.1 Number of daily accidents from 2015 to 2019 in Skudai

Table 28.2 Descriptive statistics of daily accidents in Skudai

Years	Total daily accident	Mean	Standard deviation	Kurtosis	Skewness
2015	4781	13.0986	5.3356	0.4616	1.1722
2016	5313	14.5164	5.7561	-0.5687	0.8829
2017	4963	13.5973	5.4601	0.1353	1.1217
2018	6732	18.4438	7.0411	-0.0927	0.8845
2019	6265	17.1644	6.4459	-1.2398	0.5207

above 10 days among 365 days. It can also visualize that there was a bell shape in 2015, 2017, and 2018, but there was not an asymmetrical shape in 2015 and 2017 while there was more symmetry shape in 2018. It indicated that there have more frequently occurred road accidents in 2018 in Skudai. In 2016 and 2019, the pattern of the charts will more flat than in other years.

Moreover, there has an outlier in 2015 and 2016 respectively, which was 38 cases per day of road accidents in Skudai. In 2019, there has a point far away from all of the data points, which was located at 42 cases of accidents per day as the most serious number among these five years.

Among these five years, the highest total number of daily accidents is in 2018, which is 6732 cases, but the lowest total number of daily accidents is in 2015 with 4781 cases. Since the cases of accidents are decreased in 2019, the total number of accidents still has a highest record of above 6000 cases. The mean was affected by the total number of accidents, if there have a higher number of accidents, so the mean will be affected to be higher values. The largest standard deviation is in 2018 as 7.0411 because the data points are more spread out from the mean. It is also affected by the higher value of the mean. The standard deviation in 2015 is 5.3356, which is the lowest. It indicates that the data points are closely clustered to the mean. In 2016 and 2017, the standard deviation is 5.7561 and 5.4601 respectively.

In 2015, the value of kurtosis is 0.4616 as greater than zero, which indicates that the tail of the distribution is heavy which means that the outliers are presented. Based on Fig. 28.1, the outlier in 2015 can be determined at 38 cases of daily accidents in one day. Since the kurtosis value in 2017 is 0.1353 which is closer to zero, it may assume that is a normal distribution. In 2016, 2018, and 2019 of kurtosis values, there are less than zero as negative kurtosis. This can be concluded that the tail of the distribution is light and flatter than a normal curve, which implied that there may no outlier. Since all the values of skewness from 2015 to 2019 are positive, then the data are skewed right. In 2015 and 2017, the values of skewness are greater than 1, so the data are highly skewed to the right. Meanwhile, in 2016, 2018, and 2019, the values of skewness are between 0.5 and 1. This can be concluded that the data are moderately skewed to the right in those three years. All of the year is only one day without any road accident, except in 2019. It revealed that Skudai has occurred road accidents every day in 2019.

Among 365 days in 2015, the highest number of days that occurred accidents is 38 days with 13 cases of accidents. This implied that there have 494 cases of accidents. In 2016, there have 28 days with 11 cases, 12 cases, and 14 cases of accidents as the highest number of days. Besides, the highest number of days in 2017 is 37 days with 13 accidents. In 2018 and 2019, the highest number of days that occurred accidents is 29 days at 18 cases of accidents and 25 days that have involved in 21 cases of accidents per day. Next, the highest number of accidents in one day is 38 cases in 2015 and 2016. In 2017, there have 33 cases of accidents that occurred in one day. The highest daily accident in 2018 was 39 cases while in 2019, there have 42 cases as the most serious number of accidents in one day. It can be concluded that an increasing trend of daily accidents from 2015 to 2019.

For the number of daily fatal accidents, the results of grouped bar chart showed in Fig. 28.2 to visualize the pattern of the fatal accident that occurred with the number of days in Skudai. To support the finding in the bar chart, the summary of statistics for the fatal accidents is presented in Table 28.3.

Table 28.3 shows there have eight cases of fatal accidents from the accidents in 2015. Besides, the highest total of fatal accidents is 18 cases in 2016. In 2017, there have 10 cases of total fatal accidents in Skudai area. 12 cases of fatal accidents happened in 2018. In 2019, there have 9 cases of fatal accidents within 6265 cases of total daily accidents. The mode and median are zero because most of the days with daily accidents did not have fatal accidents. Thus, it is obtained a good result for less cases of fatal accidents in Skudai area. The highest number of days that occurred fatal

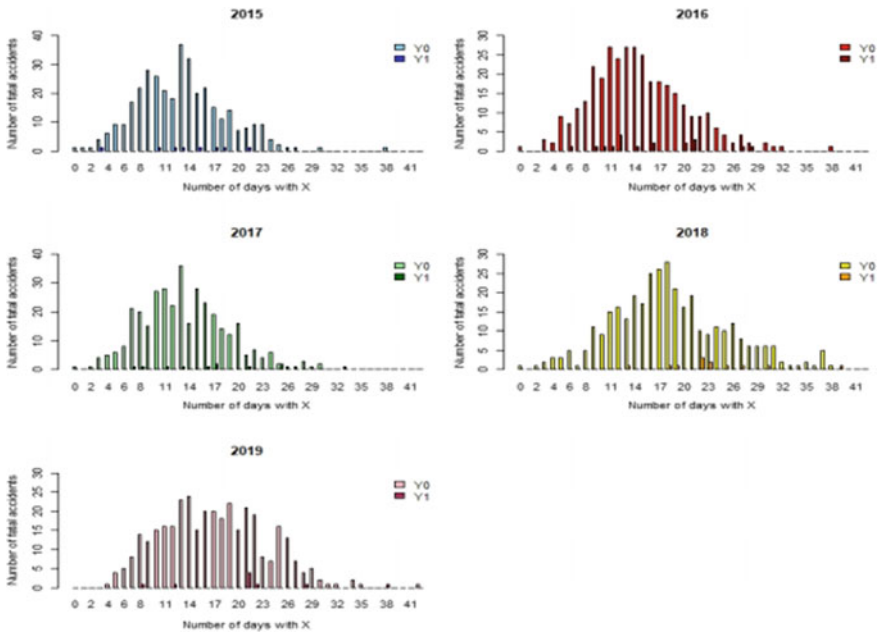


Fig. 28.2 Number of daily fatal accidents from 2015 to 2019 in Skudai

Table 28.3 Summary of daily fatal accidents in Skudai

Years	Total fatal accidents	Mode	Median	The highest number of days that occurred fatal accidents
2015	8	0	0	1
2016	18	0	0	4
2017	10	0	0	2
2018	12	0	0	3
2019	9	0	0	4

accidents in 2016 and 2019 are four days among 366 days and 365 days respectively. In 2018, there have three days occurred fatal accidents as the highest number of days. Thus, it can be concluded that the number of fatal accidents from 2016 to 2019 is reduced. It reflects those fatal accidents in Skudai have a decreasing trend to avoid fatal accidents for every road user.

The bar charts of legends for $Y = 0$ implied that is no fatal accident while $Y = 1$ implied that has fatal accidents. Figure 28.2 shows that there had fewer fatal accidents in 2015, but there have higher fatal accidents cases in 2016. The highest number of fatal accidents was four cases while the second-highest was three cases in 2016. In 2017, it can visualize that the fatal accidents were decreased. In 2018, the number of fatal accidents was increased again and the highest number of fatal accidents was three cases. There were decreased cases of fatal accidents in 2019, but the highest fatal accidents were four cases from the daily accidents. It can be indicated that a decreasing trend of fatal accidents in Skudai among these five years.

For the number of daily fatal victims, the results of grouped bar chart showed in Fig. 28.3 to visualize the pattern of the fatal victims that occurred with the number of days in Skudai. To support the finding in the bar chart, the summary of statistics for the fatal victims is presented in Table 28.4. The bar charts of legends for $Z = 0$ implied that is no fatal victim, $Z = 1$ implied that has one fatal victim and $Z = 2$ indicated that has two fatal victims in fatal accidents.

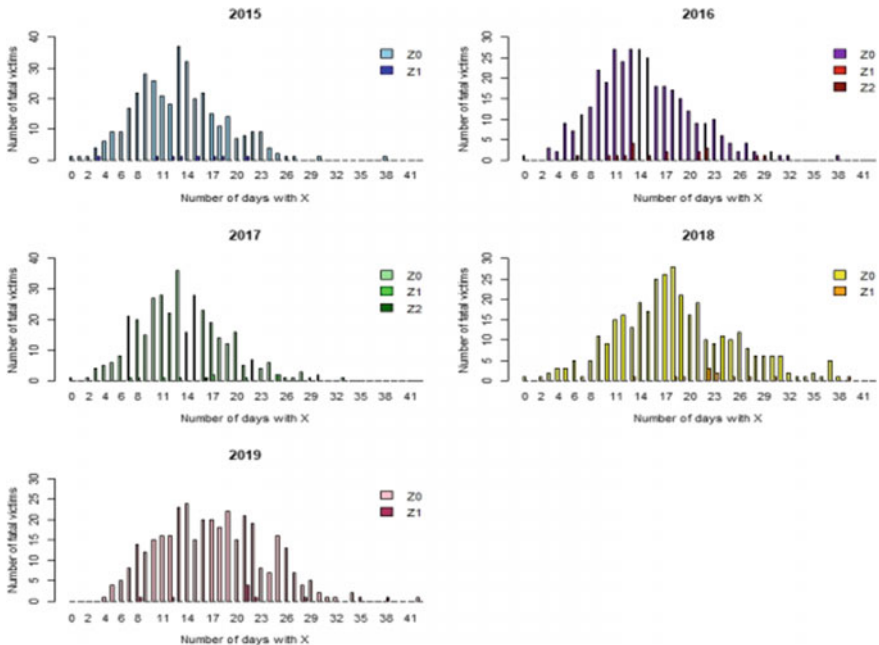


Fig. 28.3 Number of daily fatal victims from 2015 to 2019 in Skudai

Table 28.4 Summary of daily fatal victims in Skudai

Years	Total fatal victims	Mode	Median	The highest number of days that occurred fatal victims
2015	8	0	0	1
2016	19	0	0	Z = 1(4), Z = 2 (1)
2017	11	0	0	Z = 1(2), Z = 2 (1)
2018	12	0	0	3
2019	9	0	0	4

Figure 28.3 shows that there had a fewer number of fatal victims in 2015, but there have higher fatal victims who lost their lives in the road accidents in 2016. In 2017, it visualized that the fatal accidents were decreased. Based on Fig. 28.3, there have two fatal victims in the fatal accidents in 2016 and 2017. However, the fatal accidents were increased again in 2018 as well as there was decreased in 2019. It can be obtained that a decreasing trend of fatal victims in Skudai from 2015 to 2019.

Based on Table 28.4, the total number of fatal victims were involved eight people in 2015. In 2016, there have 19 fatal victims which is the highest number of daily fatal victims among the five years. Besides, the total number of fatal victims in 2017 is 11 people while in 2018 is 12 people. In 2019, there has a decreasing trend for the total of fatal victims as only nine people. Other than that, the mode and median are zero from 2015 to 2019. It is because most of the days did not have any fatal accidents, so there would not have fatal victims caused by road accidents. The highest number of days that has the fatal victim in 2015 is only one day, which means that there has only one fatal victim involved in each of the fatal accidents. In 2016 and 2017, there only one day has two fatal victims from the accidents. Besides, the highest number of days for daily fatal victims in 2018 is three days while in 2019 is four days. As a conclusion, this can be concluded that the number of fatal victims was decreased in Skudai from 2016 to 2019.

28.5.2 Parameter Estimate

According to Eq. 28.9, this study applied the method of moment to estimate $\hat{\theta}$, $\hat{\lambda}$, $\hat{\eta}$ and \hat{p} of parameters. Table 28.5 shows that the four parameters of the estimation from 2015 to 2019 in Skudai.

For the daily accident rate ($\hat{\theta}$), it fluctuated from 2015 to 2019. There is 8.8850 in 2015 to 9.6086 in 2016, which is increased, then in 2017 is 9.1828 as decreased. Besides, there is increased to 11.2496 in 2018 and decreased to 11.0321 in 2019. The overall result obtained that the daily accident rate was increased among the five years, which implied that the daily accidents in Skudai rose. For accident hazard level ($\hat{\lambda}$), the results presented the fluctuated pattern from 2015 to 2019, but the overall result showed slowly increased in Skudai. Since all the values of accident hazard

Table 28.5 Parameter estimate $\hat{\theta}$, $\hat{\lambda}$, $\hat{\eta}$ and \hat{p}

Estimators	Years				
	2015	2016	2017	2018	2019
$\hat{\theta}$	8.8850	9.6086	9.1828	11.2496	11.0321
$\hat{\lambda}$	0.3217	0.3381	0.3247	0.3901	0.3573
$\hat{\eta}$	0.0011	0.0024	0.0015	0.0011	0.0009
\hat{p}	0.0017	0.0034	0.0020	0.0018	0.0014

levels are greater than zero, it indicated that there has no effort done by road users to reduce the accidents. In 2015, the daily fatal victim rate ($\hat{\eta}$) is 0.0011 in 2015. This rate rose to 0.0024 in 2016 and dropped in 2017, which is 0.0015. After that, there is decreased to 0.0011 in 2018 and decreased to 0.0009 in 2019. Although this rate is increased in 2016, there is reduced from 2017 to 2019. It indicated that was reduced the leading cause of death of road accidents in Skudai. The probability of daily fatal accidents (\hat{p}) is increased from 2015 to 2016, which are 0.0017 and 0.0034 respectively. However, this probability dropped from 2017 as 0.0020 to 2018 as 0.0018, then decreased again to 0.0014 in 2019. Since the probability of daily fatal accidents is decreased followed by years. Thus, this is implied that the road users have paid attention and responsible when using the vehicles on the roads to avoid fatal accidents.

28.5.3 Goodness of Fit Test

Three stochastic models have applied to investigate whether the model is significantly fit to the observed data. Based on the GP model, the model used the number of daily accidents as observed. The result shows only 2017 and 2019 are significantly fit to the GP model. However, the GP model is not appropriate for the observed data in 2015, 2016, and 2018. Hence, this condition did not have strong evidence to obtain all of the years are fit to the GP model. For the GPQB model, data of the number of daily fatal accidents are used as the observed data. The results present that all of the p -values were less than 0.05 as 0.000. The test statistics values of the goodness-of-fit test are very high because there have many zero data within the years when there are no fatal accidents frequently happened. There is sufficient evidence to reject the null hypothesis. Therefore, it concludes that the GPQB model is unsuitable to fit the data of fatal accidents from 2015 to 2019.

Meanwhile, the GPGP model applied to the number of daily fatal victim datasets as the observed data. The results showed that there only one year that the GPGP model is fitted significantly to the data in 2017. Furthermore, the other years of outcomes failed to fit with the GPGP model. It implied that the GPGP model is not fitted significantly to the data in 2015, 2016, 2018, and 2019.

28.6 Conclusion

The aim and the objectives are achieved in this research. The trend and pattern for daily traffic accidents were visualized and shows some meaningful behaviours. The findings from this research might help the authority in traffic management to take the necessary actions for reducing the number of accidents and improve the quality of the traffic based on the condition at Skudai.

In this study, the overall road accidents are increased from 2015 to 2019 in Skudai. It also revealed that the highest case in 2019 is 42 cases of accidents in one day. This finding resulted that the road accidents frequently occurred in Skudai recently. In contrast, the total number of fatal accidents is dropped to less than ten cases. It shows a good result for the reduction of the leading death for road accidents in Skudai from 2015 to 2019. For parameter estimation, the overall accident rate is increased from 2015 to 2019. It was obtained that the average of daily accidents is increased in Skudai. The accident-hazard level is slowly increased, which indicated that the road users no effort to reduce or avoid road accidents in five years as the accident hazard level values are greater than zero. However, the probability of fatal accidents and daily fatal victim rates are decreased because of the lower trend of fatal accidents and only fewer people involved. It also obtained those drivers had emphasized safety awareness of death when driving on the road. The finding of the goodness of fit test presented the GP model is significantly fitted with 2017 and 2019 of daily accident data while the GPGP model is only fitted with 2017 of daily fatal victim data. For the GPQB model, the result obtained that all of the variables are not fitted with the model. This may reflect that the GPQB model is unsuitably applied in the latest fatal accident data due to the less fatal accidents in Skudai.

Acknowledgements This research was supported by Universiti Tun Hussein Onn Malaysia (UTHM) through Tier 1 (vot H785). The authors sincerely appreciate and grateful to the officers at Ibu Pejabat Polis Diraja Bukit Aman and Ibu Pejabat Polis Daerah Johor Bahru (Utara), Skudai for their enthusiasm in providing field data during the Movement Control Order (MCO) 01. The authors are also extremely thankful to the reviewers for their beautiful remarks.

References

1. A.B. Labana, V.A. Parik, V.P. Parekh, Literature review on road accident analysis a case study on Dahod to Jhalod section of N. H. 113. *Int. J. Sci. Res. Dev.* **3**(2), 2338–2341 (2015)
2. A.A. Mustafa, Database development of road traffic accident case study Johor Bahru, Malaysia. *J. Soc. Transp. Traffic Stud.* **3**(1), 1–8 (2015)
3. Ministry of Transport Malaysia, *Transport Statistics Malaysia 2018* (Official Portal Ministry of Transport Malaysia, Malaysia, 2018)
4. M.A. Djauhari, Stochastic pattern of traffic accidents in Bandung city. *IATSS Res.* **26**(2), 85–91 (2002)
5. C.B. Thompson, Descriptive data analysis. *Air Med. J.* **28**(2), 56–59 (2009)
6. S. Chaudhari, *Descriptive Statistics* (2018). <https://towardsdatascience.com/descriptive-statistics-f2beeaf7a8df>

7. A. Rapoport, in *Stochastic Modelling of Social Processes*, ed. by A. Diekmann, P. Mitter New Aspects of Stochastic Model Building in the Social Sciences (Academic Press, Inc, London 1984)
8. H.M. Taylor, S. Karlin, *An Introduction to Stochastic Modelling*, 3rd edn. (Academic Press, California, 1998)
9. P.C. Consul, G.C. Jain, On some interesting properties of the generalized poisson distribution. *Biom. Z.* **15**(7), 495–500 (1973). <https://doi.org/10.1002/bimj.197301150707>
10. B.R. Bhat, *Modern Probability Theory*, 3rd ed. (New Age International (P) Limited, Delhi, 1999)
11. R.E. Leiter, M.A. Hamdan, Some bivariate probability models applicable to traffic accidents and fatalities. *Int. Stat. Rev.* **41**(1), 87–100 (1973). <https://doi.org/10.2307/1402790>
12. R. Shanmungan, J. Singh, Some bivariate probability models applicable to traffic accidents and fatalities. *Statist. Distrib. Sci. Work* **6**, 95–103 (1981)

Chapter 29

The FGVH and GDEX Stock Price Modeling Using ARIMA and Holt's Linear Trend Methods



Norhaidah Mohd Asrah , Rabia'tul Adawiyah Sua'aif,
and Nur Hazlina Abdul Wahab

Abstract Nowadays, stock market represents a company performance and an essential asset to promote an investment to the company. It is crucial to any company for the reflection on the growth and fall of the company. It helps the company to attract more investors when the stock is growing and can contribute to the growth of an economical state of a country. The analysis of the stock market price would be essential to know the fluctuation of the stock market prices. When the company's capital is increase, it will be adequate for the growth of the company. Thus, it will attract more investors to rise the economic trade, growth, and prosperity of the company. To boost the confidence of investors to invest, it is important to study the stock price market and find the suitable model so that the direction of the stock price can be determine. Hence, this study focuses on two different type of industries stock market namely Felda Global Ventures Holdings (FGVH) and GD Express Carrier Bhd (GDEX). The plantation industry is represented by the FGVH stock market, while the services industry is represented by the GDEX. The modeling of both stock price is studied using the Autoregressive Model Integrated Moving Average Model (ARIMA) and the model of exponential smoothing, Holt's linear trend. The results revealed that the best model for FGVH stock market is ARIMA (0, 1, 2) and for GDEX stock market is the Holt's linear trend with smoothing constant.

N. M. Asrah (✉)

Data Analytics, Sciences and Modeling Focus Group, Universiti Tun Hussein Onn Malaysia,
84600 Pagoh, Johor, Malaysia
e-mail: norhaida@uthm.edu.my

N. M. Asrah · R. A. Sua'aif · N. H. A. Wahab

Department of Mathematics and Statistics, Faculty of Applied Sciences and Technology,
Universiti Tun Hussein Onn Malaysia, 84600 Pagoh, Johor, Malaysia

29.1 Introduction

29.1.1 *Background of the Study*

The stock market is a set of markets and exchanges where regular operations such as buying, selling, and issuance of publicly traded company shares take place [1]. Businesses could use stock trading to raise funds to pay off debt, introduce new goods, and expand operations. During rising or bull markets, businesses undertake capital investments, which may lead to greater consumer spending. Increased consumer spending may lead to economic growth since consumer spending in stock markets is an important component of the gross domestic product [2]. Every stock price has its own behavior and to estimate the stock price behavior that valuable in the market and contributed to the future, there are a few theories that must understand [3]. For instance, stock market performance is dependent on the fundamental macroeconomic factors and a predictor for the expected performance [4]. Predicting the stock price is critical since the stock price tends to reflect the state of the stock market, and the stock market plays a critical role in boosting economic growth. Furthermore, by correctly predicting the stock price, investors and corporations can invest at the perfect time and benefit [1]. The use of time series analysis can aid in the prediction of future occurrences, particularly in the commercial world. One of the major events that will influence investors and corporations in the future is the time series of stock price data.

Hence, this study focusses on the stock price of two different type of industries and companies. The first stock price refers to the performance of the largest company that produces palm oil as the main production known as Felda Global Venture (FGVH) Holding Berhad [5]. Being the one of the largest plantations' companies in Malaysia, FGVH stock price forecasting is important to present the company market performance. The stock price able to help a particular company to predict the financial market analysis and usually used to determine the value of company stock for upcoming days [6]. There are three main sectors in the FGVH, known as plantation, sugar, and logistics. These sectors allowed greater convergence and open more opportunities for collaboration and innovation in the future. Besides being the one of the crude palm oil producers, the FGVH company also leading the refined sugar producer in Malaysia with 585 domestic market share and over 2.25 million tons annual capacity production of refined sugar [7]. Based on this reputation, the FGVH stock price is important to present the company performance in the plantation sector market. However, there is non-consistency and declining in the FGVH stock price starting from 2013 until 2020 due to the several issues [6]. Therefore, it is crucial to study the FGVH stock price performance to understand the behaviors and the value of the stock price in the future.

The second company is GD Express Carrier Berhad (GDEX) [8]. This company is active in the provision of express delivery, transport, rent of electronic devices to relevant businesses, maintenance of facilities and properties, insurance, and logistics services. In Malaysia, GDEX has a large network coverage. GDEX provides a variety

of solutions to satisfy the needs of its customers. GDEX made a pre-tax profit of RM1.1 million in 2003, and it is currently preparing to list on the Malaysian Exchange of Securities Dealing and Automated Quotation (MESDAQ). After that, in 2005 GDEX stock was successfully listed on Ace Market of Bursa Malaysia [9]. However, in 2013 the GDEX stock code (0078) was planned to introduce the dividend policy and seek the overseas development opportunities [10]. Nowadays, the investors are confident to invest in GDEX stock price because of the high growth trend and the cooperation with the international companies such as the Chinese giant internet, the Alibaba Group Holding Ltd [9].

29.1.2 The Objectives of the Study

The objectives of this study are to analyze the FGVH and GDEX stock price by using the ARIMA and Holt's linear trend methods. When there is limited information about the underlying data generation process or when there is no appropriate explanatory model that ties the prediction variable to other explanatory factors, this modelling method is highly beneficial [11]. The main discussion for this study is referred to the studies from [5, 8]. Both studies have been conducted previously and yielded an interesting result. It took our interest to reviewed and discussed both. The limitation of this study is the analysis only focus on the FGVH and GDEX stock price only. Based on these results, perhaps this study can provide more information whether ARIMA and Holt's linear trend are suitable to model different type of industries stock price.

29.2 Methodology

29.2.1 Data Collection

The dataset for both stock price is collected from <https://investing.com>. The FGVH stock price data are collected from 2nd January 2015 until 13th July 2020. Meanwhile, the GDEX stock price data are collected from 2nd May 2017 until 31st March 2020.

29.2.2 The Box-Jenkins Method

In the preliminary study, the Box-Cox transformation is needed to transform a dependent variable into normal shape and act as a logarithm that helps to stabilize the time-series variance. The transformation from a non-stationary series to stationary

in mean and variance is important in the Box-Jenkins method. The variance is considered not stable when the value of λ is not equivalent to 1. The value λ is known as the transformation parameter [12]. Then, the stationarity of time series pattern will give unpredictable patterns in the long- term period. Therefore, it is important to ensure the stationarity of the time series pattern [13]. The ACF plot can be useful to identifying the non-stationary time series. This step is the main step for identification and parameter selection in time series model.

The unit root test as Augmented Dickey-Fuller (ADF) and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) can be used to check the stationarity. If the dataset is not stationary, hence the differencing is required. Differencing is a repeated process until the stationary time series is present. When there is no differencing is involved, the abbreviation ARMA will be used. The autoregressive model or usually called as AR model in Autoregressive Integrated Moving Average (ARIMA) and an autoregressive model of order p can be written as in (29.1):

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + e_t \tag{29.1}$$

where y_t are the original series, ϕ 's called are the autoregressive parameters to be estimated, c is the average of the changes between consecutive observations and e_t refer to white noise. Moving average model used the past forecast errors in a regression-like model or other words, MA model measure the adaptation of new forecasts to the prior forecast errors [14]. Thus, the moving average model of order q can be written as in (29.2):

$$y_t = c + e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \dots + \theta_q e_{t-q} \tag{29.2}$$

In ARIMA, model identification is required to choose the value of p, d and q where refer to the autoregressive order, differencing and moving average order respectively in ARIMA (p, d, q) model. The Autocorrelation Function and Partial Autocorrelation Function (ACF and PACF) plots are used to estimates the p, d and q values. Using both plots, the suitable values for order p and q for both models, AR and MA can be identified adequately. In ARIMA, the term “integrated” refers to the inverse of differencing. Equation (29.3) can be written as the first model:

$$y'_t = c + \phi_1 y'_{t-1} + \dots + \phi_p y'_{t-p} + \phi_1 e_{t-1} + \dots + \phi_q e_{t-q} + e_t \tag{29.3}$$

where y'_t represent the differenced series and differencing may have been more than one time. The predictors on the right side include both lagged y_t values and lagged errors. Combining the model may be intricate and challenging, but working using backshift notation, as illustrated in (29.4), makes it easier:

$$(1 - \phi_1 B - \dots - \phi_p B^p)(1 - B)^d y_t = c + (1 + \phi_1 B + \dots + \phi_q B^q) e_t \tag{29.4}$$

Table 29.1 The model identification using ACF and PACF patterns

Model	ACF patterns	PACF patterns
MA(q)	Cut off after lag q	Dies down or exponentially decaying
AR(p)	Dies down or exponentially decaying	Cut off after lag p

The right-hand side refer to the MA(q) while left-hand side indicates the AR(p) and d differencing, respectively. The parameter estimation is conducted due to determine the possible appropriate model by involving the statistical tools such as Autocorrelation Function (ACF) and Partial Autocorrelation Functions (PACF). Table 29.1 show a guideline to identify the significant model using ACF and PACF patterns.

The most appropriate model is estimate and selected due to the lowest value of Akaike’s Information Criterion (AIC), corrected Akaike’s Information Criterion (AICc) and Bayesian Information Criterion (BIC). The complementary approaches represented the residual analysis from the fitted model and equation of residual can be written as in (29.5):

$$Residual = Actual\ value - Predicted\ value \tag{29.5}$$

The residuals value should close to the white noise properties, due to the significant model and the parameter estimates are close to the actual values. Besides, autocorrelation plot will be applied to see the further structure of the large correlation value can be found. The small autocorrelation (ACF) and partial autocorrelation (PACF) plot of residuals, thus the model will be considered as adequate model and forecasts will be conducted.

29.2.3 The Holt’s Linear Trend Method

Holt’s linear trend method also known as Holt’s trend corrected exponential smoothing, is one of the methods of exponential smoothing. It is used when the time series is increasing or decreasing approximately (with a trend) at a fixed rate [14]. This model involves a forecast equation and two smoothing equations which are α and β . According to (29.6), it explains that if the parameter β_0 and β_1 are slowly changing over time, therefore, regression will be used to future value of y_t .

$$y_t = \beta_0 + \beta_1 t + \varepsilon_t \tag{29.6}$$

There are two estimates, the level estimate and trend estimate as in (29.7) and (29.8):

$$l_t = \alpha y_t (1 + \alpha)(l_{t-1} + b_{t-1}) \tag{29.7}$$

$$b_t = \beta(l_t - l_{t-1}) + (1 - \beta)b_{t-1} \quad (29.8)$$

where l_t represent an estimate of the level series at time t , α is a smoothing constant for the level ($0 \leq \alpha \leq 1$), b_t denotes an estimate of the trend which is the slope of the series at time t , followed by β indicates the smoothing constant for the trend ($0 \leq \beta \leq 1$). The level estimate l_{t-1} is the estimate of the time series constructed in period $t - 1$ called as the permanent component. The trend estimate, b_{t-1} which is the estimates of the growth rate of the time series constructed in period $t - 1$, called as trend component.

$$\hat{y}_{+p}(T) = (l_T + pb_T), (p = 1, 2, 3, \dots) \quad (29.9)$$

The point forecast made at time t for y_{T+p} as in (29.9), where p -step-ahead forecast is corresponding to the last value estimated level plus p times last estimated trend value. Therefore, the forecast values for the next period are the linear function of p .

29.3 Results and Discussion

29.3.1 The Time Series Plot

The time series plot of FGVH stock price data is as Fig. 29.1a and the GDEX plot as in (b). According to Fig. 29.1a, the FGVH stock price showed a downward cyclic movement. It decreased drastically for the first 500 days, from RM4.00 to RM2.00. Then, the stock has stable price between RM1.00 to RM2.00 for next 1000 days. Then, the stock price drops to less than RM1.00. Meanwhile, the time series plot for GDEX stock price in (b), it also shows a downward trend and seasonality pattern. The stock price decreased from the beginning and reached the lowest price which is RM0.205 in March 2020.

The result for normality test of FGVH stock price is not normal since Fig. 29.2a show the p -value is less than 0.005. Thus, the square root transformation of Box-Cox transformation was conducted to stabilize the variance. Based on the Fig. 29.2b, the optimal lambda is equal to 1 which indicate that the data is normal and the transformation data equivalent to original data. At the same time, the stock price for GDEX is also not normal as in Fig. 29.2c. After the square root transformation of Box-Cox, noticed that the optimal lambda is equal to 1 as in Fig. 29.2d and no further transformation is required.

Next, the stationarity for both stock price is analyzed and need differencing since both stock price is not stationary. The ACF and PACF plots for both stock price after differencing can be referred to Fig. 29.3. The autocorrelation function (ACF) and partial autocorrelation function (PACF) then are used to estimate the parameters of ARIMA. ACF and PACF plot used to indicate the appropriate value for p and q . The data may follow an ARIMA ($p, d, 0$) model or AR model if ACF and PACF plots

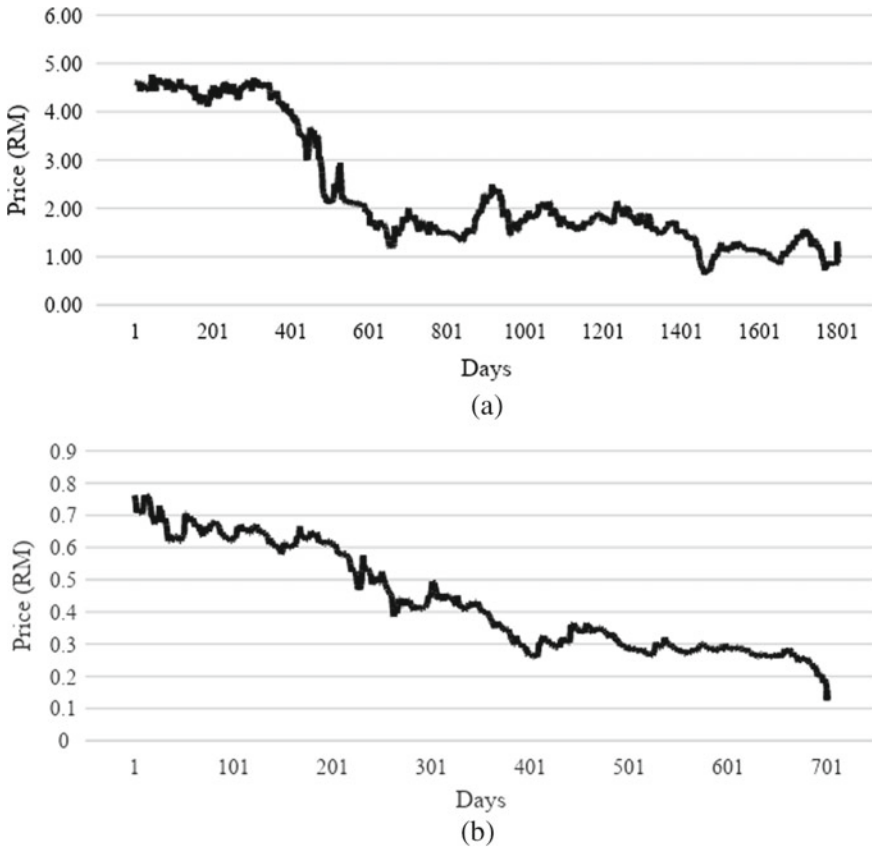


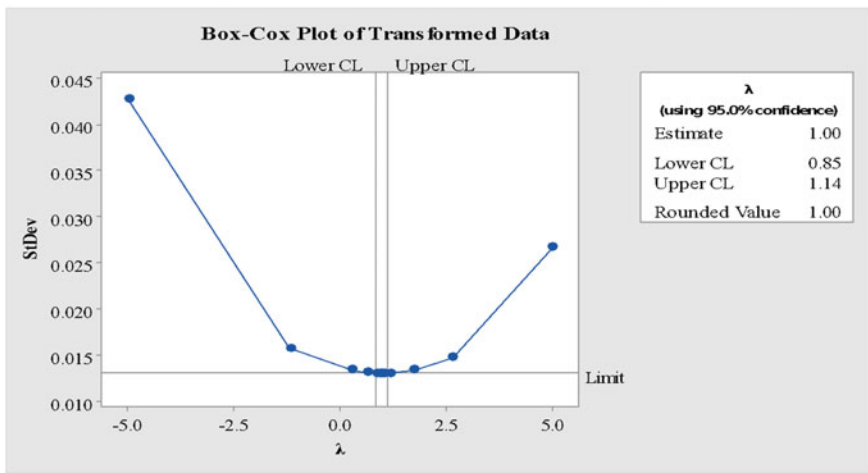
Fig. 29.1 The time series plot of FGVH stock price (a) and GDEX stock price (b)

is exponentially decaying or sinusoidal. There is also a significant spike at lag p but none beyond the lag p . While for ARIMA $(0, d, q)$ model or MA model if the PACF plot is showing the exponentially decaying or sinusoidal pattern and there is also a significant spike at lag q but none beyond lag q .

Figure 29.3a for FGVH stock price shows that there is one small spike in the ACF plot at lag 2 with the value of lag with 0.061 and then no significance spikes thereafter. While in the PACF plot in Fig. 29.3b, it showed the sinusoidal pattern where there is one spike with the lags and another two spikes at lag 20 and 21 and then no significant spikes thereafter. The spike that located far from the first spike can be ignore in each plot if it just outside the limits but not in the first lags. After all, the probability of a spike being significant by chance is about 30 but 32 spikes in each plot were plotting. The pattern in the second spike is what would expect from ARIMA $(0, 1, 2)$ as the PACF plot tends to decay exponentially and there is significant at lag 2 in ACF but not beyond the lag 2. Besides, ARIMA $(0, 1, 2)$ model, ARIMA $(0, 1, 3)$ model and ARIMA $(2, 1, 2)$ also have probability to be the appropriate model.



(a)

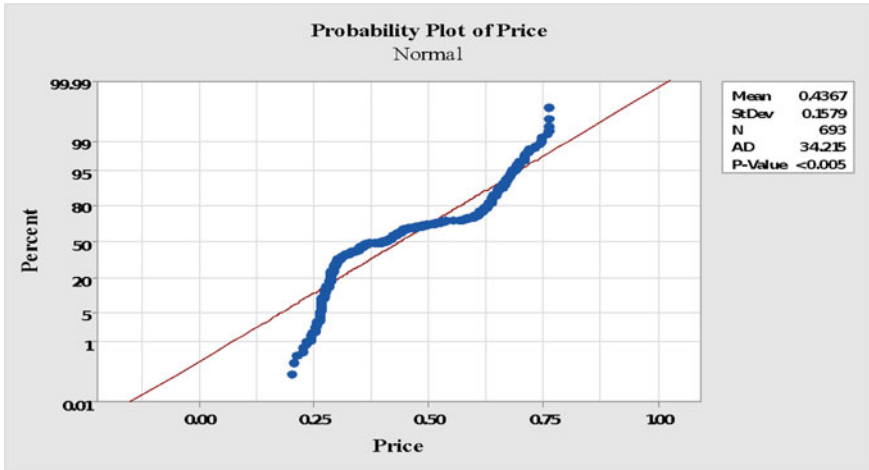


(b)

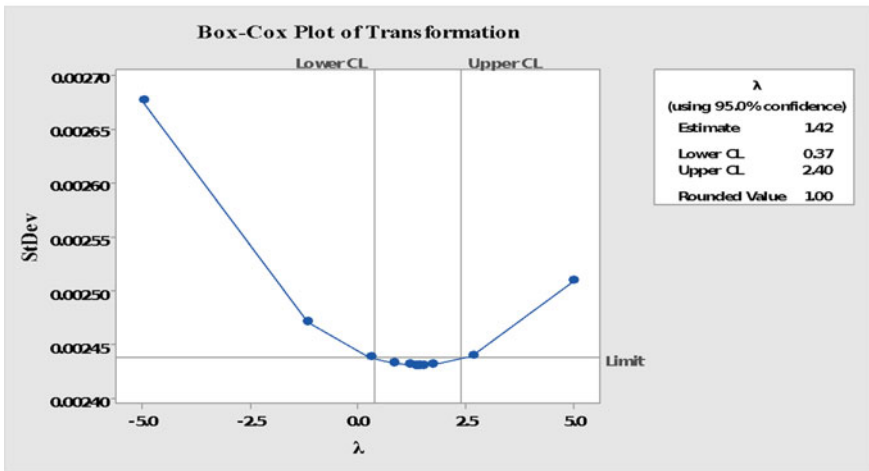
Fig. 29.2 The probability plot and the Box-Cox transformation for FGVH stock price in (a, b) and GDEX stock price in (c, d)

In Table 29.2, the model ARIMA (0, 1, 2) shown the minimum value of Akaike’s Information Criterion (AIC), corrected Akaike’s Information Criterion (AICc) and Bayesian Information Criterion (BIC) if compared to another two models. These results obtained by using R. Therefore, ARIMA (0, 1, 2) model will be chosen as the best model to modeling the stock price of FGVH.

For the GDEX stock price in Fig. 29.4c, d, the value of d is 1, because of the differencing. These leave the value of p and q to be represented as either the value of 3 or 5. Therefore this stock price has two different model which are (5, 1, 5) and (3, 1, 3). The results of each model for GDEX stock price are obtained by using



(c)



(d)

Fig. 29.2 (continued)

Minitab 18, which includes the final estimates of parameters, residual sum of squares, modified Box-Pierce (Ljung-Box) Chi-square statistic and the ACF and PACF plot of the residuals for the dataset.

Based on Table 29.3, the final estimates of parameters for the (5, 1, 5) model, both autoregressive term and moving average term, has p -value of 0.000 and 0.001 and smaller than the significance level of 0.05. Therefore, the autoregressive term coefficient is statistically significant, and it should be retained in the model. Next, the (3, 1, 3) model has a p -value of 0.0000 for both autoregressive and moving average terms. Therefore, the autoregressive term coefficient is statistically significant, and

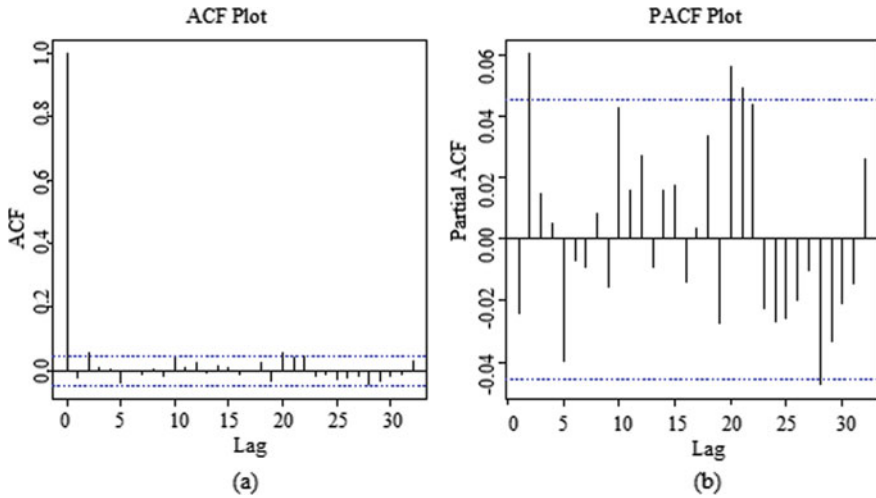


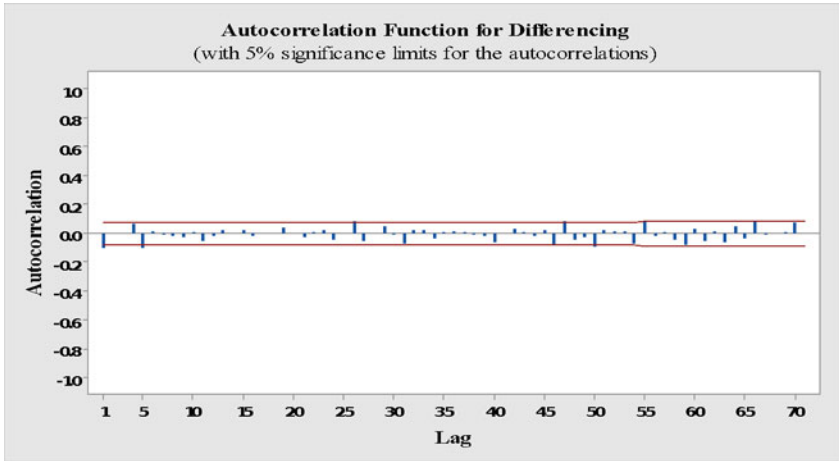
Fig. 29.3 The ACF and PACF plot for FGVH stock price as in (a, b)

Table 29.2 The ARIMA models for FGVH stock price

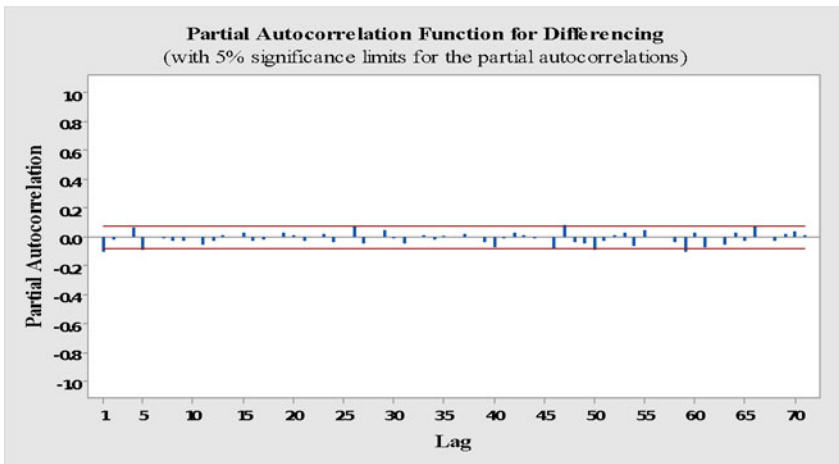
Model	Akaike’s information criterion (AIC)	Corrected Akaike’s information criterion (AIC _c)	Bayesian information criterion (BIC)
ARIMA (0, 1, 2)	−9049.76	−9049.75	−9033.18
ARIMA (0, 1, 3)	−9048.41	−9048.39	−9026.3
ARIMA (2, 1, 2)	−9046.94	−9046.91	−9019.3

it should be retained in the model. Then models are compared by using the residual sums of squares. The smaller values of the sums of squares and the mean square indicate a better fitting model. Both models of (5, 1, 5) and (3, 1, 3) have a mean square of 0.0000156 and 0.0000158 respectively. Therefore, the (5, 1, 5) model shows the lowest value and it will be chosen as the best model as well. The comparison by the Ljung-box Chi-square statistics used *p*-value to be compared to the significance level of 0.05. If the *p*-value of a model is higher than the level of significance, then the residuals of the model are independent. The (5, 1, 5) model is the most suitable model to be chosen as the best model among the other models.

Meanwhile, the appropriate smoothing parameter for Holt’s linear trend model is as Table 29.4. The level indicated by alpha for FGVH stock price is 0.975 while the trend that indicated by beta is 1.000, with the initial state of the level was equal to 2.472 and −0.1282 for the initial trend. As for GDEX stock price, the smoothing parameter for alpha is 0.9 and beta is 0.



(a)



(b)

Fig. 29.4 The ACF and PACF plot for GDEX stock price as in (a, b)

29.4 Conclusions and Recommendations

This study applied the ARIMA and Holt’s linear trend methods to analyse the time series dataset for FGVH and GDEX stock price in Malaysia. The FGVH stock price represent the plantation company while the GDEX stock price represent the courier service company. Both stock price is modeled by using both methods approach. Both stock price has downward trend with cyclic movement for FGVH stock price and seasonality for GDEX stock price. Both stock prices are not normal at the beginning

Table 29.3 The ARIMA models for GDEX stock price

Model	Final estimates of parameter (<i>p</i> -value)	Mean of square	Ljung-Box Chi-square (<i>p</i> -value)
ARIMA (5,1,5)	AR (0.0000)	0.0000156	
Lag 12	MA (0.0010)		0.441
Lag 24			0.975
Lag 36			0.944
Lag 48			0.766
ARIMA (3,1,3)	AR (0.0000)	0.0000158	
Lag 12	MA (0.0000)		0.1
Lag 24			0.697
Lag 36			0.654
Lag 48			0.556

Table 29.4 The Holt’s linear trend model

Stock price	Smoothing parameters	
	Alpha	Beta
FGVH	0.975	1
GDEX	0.9	0

and need to transform using the square root transformation. Then, both stocks are check for the stationarity and need differencing.

Based on the results, the best model for FGVH stock price is ARIMA (0, 1, 2) and the Holt’s linear trend with alpha 0.9 and beta 0, is the best model for GDEX stock price. Although both stocks are analyse using ARIMA and Holt’s linear trend, it shows that ARIMA is more suitable for FGVH stock price and Holt’s linear trend is more appropriate for GDEX stock price. For GDEX stock price, the modeling is relatively more effective because it does not have after-effect. However, there are many factors such as regional and global economic condition, socio-political conditions, poor-corporate governance, varying policies of government and psychological factors of investors that will affect the fluctuation of stock price in real situation.

It is recommended that for the future study, more stock on plantations can be includes such as Sime Darby Plantation Bhd (SIPL), Genting Plantation Bhd (GENP), TH Plantation Bhd (THPB), Hap Seng Plantations Holdings Bhd (HAPP) and BLD Plantation Bhd (BLDN). At the same time, more stock in carrier services such as Pos Malaysia & Services Holding Bhd (PSHL) and Nationwide Express Holdings Bhd (NEXP) can be includes too. Perhaps that the reason why ARIMA model is suitable for FGVH, and Holt’s linear trend is more appropriate for GDEX can be answered. Other than that, the range time of each set data can be improved by using the same range of time.

Acknowledgements This research was made possible by funding from research grant Tier 1 (vot H846) provided by Universiti Tun Hussein Onn Malaysia. The authors would like to thank the Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia for the facilities provided that make the research possible.

References

1. J. Chen, *Investment Analysis: The key to Sound Portfolio Management Strategy* (2019)
2. C. Basu, *What is a Stock and How Do Stocks Affect the Economy?* <https://finance.zacks.com/stock-stocks-affect-economy-2233.html>
3. F. Ruhani, S. Lecturer, Review of the literatures on stock price behavior of Malaysia. *Int. J. Islamic Bus. Manag.* **2**(2), 32–38 (2018). <https://doi.org/10.46281/ijibm.v2i2.219>
4. M. Masih, M.K. Nor, *Do Spot and Future Palm Oil Prices Influence the Stock Market Prices of a Major Palm Oil Producer? The Malaysian experience* (2016). <https://doi.org/10.13140/RG.2.1.3253.7368>
5. R.A. Sua' aif, *Forecasting on FGV Holdings Bhd Stock Price using ARIMA and Holt's Linear Trend Model*, Bachelor's Degree Project, Universiti Tun Hussein Onn Malaysia (2021)
6. X.Q. Sun, H.W. Shen, X.Q. Cheng, Trading network predicts stock price. *Sci. Rep.* **4**. (2014). <https://doi.org/10.1038/srep03711>
7. FGV Holding Berhad (800165-P) *Company Overview FGV is One of the World's Largest Producers of Crude Palm Oil (CPO), Accounting for about 15% of Malaysia's Total Annual Production of CPO.* <https://www.fgvholdings.com/about-fgv/company-overview/>
8. N.H.A. Wahab, *Forecasting of GDEX Stock Price in Malaysia using ARIMA and Holt's Linear Trend Model* (Universiti Tun Hussein Onn Malaysia, Bachelor's Degree Project, 2021)
9. E. Yeong, *GDEX Tightens Collaboration with Yamato.* *The Sun Daily.*: <https://www.thesundaily.my/archive/1708651-BSARCH351610> (2016)
10. GDEX Seizes Opportunity to enter Indonesian Market. <https://www.thestar.com.my/business/business-news/2018/10/06/gdex-seizes-opportunity-to-enter-indonesian-market>
11. D. Zhang, Study on forecasting the stock market trend based on stochastic analysis method. *Int. J. Bus. Manag.* **4**(6) (2003)
12. J.W. Osborne, Improving your data transformations: applying the Box-Cox transformation. *Pract. Assess. Res. Eval.* **15**(12), 1–9 (2010)
13. F.K. Arya, L. Zhang, Time series analysis of water quality parameters at Stillaguamish River using order series method. *Stoch. Env. Res. Risk Assess.* **29**(1), 227–239 (2015)
14. R. Hyndman, G. Athanasopoulos, *Forecasting Principle and Practice.* Print Edition, August 2014, otexts.com (2014)

Chapter 30

Deep Learning Approach for Football Match Classification of English Premier League (EPL) Based on Full-Time Results



Muhaimin Muszaidi, Aida Binti Mustapha, Shuhaida Ismail,
and Nazim Razali

Abstract The trend of modeling the football match classification has become increasingly popular in the last few years, thus many classification models have been proposed with the point of evaluating the attributes that lead a football team to win, draw or lose a certain match. There are two types of approaches has been considered for classification of football matches results, which include the statistical approaches and deep learning approaches. This paper proposes a Multilayer Perceptron (MLP) and Dense Neural Network (DNN) to evaluate the performance of the classification football matches results in the terms of home win (H), away win (A), and draw (D) for games under the English Premier League (EPL). The experiment revealed that MLP produced better classification accuracy rate of 78.42% as compared to 67.63% by the standard DNN. Nonetheless, the performance of DNN can be further improved by means of hyperparameter tuning, whereby DNN achieved the highest accuracy of 70.53% when modeled in a three-dense layers with the size of 16 nodes and trained with 200 epochs.

30.1 Introduction

Football matches results has always been an interesting matter to be discuss or analyses for both data scientist and the general public. By implementing machine learning algorithm on collected data, football matches outcome can be predicted using modern data analysis techniques, combined with powerful computing resources. To date, traditional approaches for predicting the results of professional football matches use

M. Muszaidi · A. B. Mustapha (✉) · S. Ismail
Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, 84600
Panchor, Johor, Malaysia
e-mail: aidam@uthm.edu.my

N. Razali
Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn Malaysia,
84600 Parit Raja, Batu Pahat, Johor, Malaysia

the number of goals scored by each team as a baseline for measuring a team's performance and predicting possible outcome. Professional football seems to be the sport of choice in this type of study because to its worldwide appeal and the massive annual turnover of the betting markets linked with it [1]. As a result, educational contributions to this field are appealing to a large number of people and can have a considerable influence on a multibillion-dollar sector.

The approaches for anticipating match results in professional football are still very undeveloped in the academic literature, despite the significant research on the subject. Past study in this sector such as by [2] studied match outcome prediction from the aspect of betting odds system. However, hardly any of the statistical models in the literature have been able to consistently beat markets [1]. Over the years, several strategies for predicting a football match have been offered. In order to anticipate football matches, some approaches apply just quantitative data, others use just qualitative data, while yet others apply a combination of both mixed method data. A quantitative dataset contains statistical information such as head-to-head results, current accomplishments, or shots on target, but a qualitative dataset contains expert information about team and player effectiveness, player accessibility, or public criticism. Even though these approaches all suggest one technique to predict the future of a football match, comparing the many approaches that may be used to do so is a less preferred response to football match prediction of past research papers.

There are three primary challenges to developing highly accurate match outcome prediction models in sports analytics, especially in forecasting match outcomes which are data availability and quality, model assumptions and testing different models and parameters. Data availability and quality is a finding of public database of football data with the necessary statistical depth to create predicted objectives metrics is a key component of the project. However, the main football data suppliers do not make their information freely available. There is a need to examine different public football databases to locate one that is acceptable for us to use. If a suitable database is not available, web scraping techniques might be used to collect data and form a dataset.

The key goal of this study is to classify the match outcome result based on the full time match result. This study will be located in the scope of English Premier League (football). The data for this study was obtained from the website <http://www.football-data.co.uk/>, which contains a wide range of information; however, in order to keep the focus narrow, only information regarding EPL football was chosen. This dataset contains data from the English Premier League's previous ten seasons, as well as the current season, as well as statistics such as final and half-time scores, corners, yellow or red cards. As such, this paper attempt to classify full-time results data based on two function-based classification algorithms in WEKA, which are the Multilayer Perceptron (MLP) and Dense Neural Network (DNN). Finally, the results will be compared based on their accuracy, precision, recall, and F-measure along with additional comparative experiments for hyperparameter tuning.

The remainder of this paper is organized as follows. A summary of previous work on football predictions by using machine learning or deep learning is presented in Sect. 30.2. In Sect. 30.3, the experiments including the dataset, the Multilayer

Perceptron (MLP) and Dense Neural Network (DNN) are detailed out. Section 30.4 present the analysis of the comparative results and finally Sect. 30.5 concludes the paper.

30.2 Related Work

In general, football match outcomes can be described in terms of results, goals scored, goal differential, or chance of winning [3]. Statistical tools can thus be used to identify determinants and forecast football results. In some cases, the outcome variable of interest is ball possession [4]. Other studies use multivariate analysis, such as principle components analysis and clustering, to summarize many game-related variables in order to separate the best clubs [5] or to assess the players' market worth in order to discover how much a team's reputation or individual abilities influence market values [6].

Deep learning, a contemporary machine learning methodology, improves learning in hierarchical structures by utilizing deep classification architectures. Deep learning model's human intellect for processing natural signals, attracting the attention of numerous academics. Deep learning is also utilized in a variety of industries to handle massive amounts of data such as Apple, Google, and Facebook to analyze massive amounts of data [7]. Google also uses deep learning algorithms to process large amounts of data from Google Translate, Google Street View, and Google Image Search [8]. Deep learning works in football prediction includes the work by [9].

Artificial Neural Networks are computer programs that attempt to imitate the behavior of a biological neural network by containing components that are linked together that turn a transition of a collection of inputs into a desired outcome [10] was one of the first to use ANN to forecast the outcome of sporting events. He gathered data on the number of yards gained, the number of rushing yards gained, the turnover margin, the length of possession, and the betting line odds from the National Sports League's initial eight rounds (NFL). Purucker's work was later followed by [11] where he introduced new elements including rushing yards and overall yardage difference, Away team indication, turnover margin, and home team indication, among others. The problem was broken down into two categories: the result of the away team and the result of the home team which was (-1 for a loss, +1 for a victory). He used the data from 208 matches and eventually obtained a 75% accuracy rate. As a result, the outcomes were compared to eight ESPN sportscasters' forecasts. The domain experts accurately predicted 63% of the matches on average.

Work by [12] used data from 110 English Premier League matches from the 2014 to 2015 seasons as a source of information for an ANN system. They made advantage of the following features: Home and Away goals (GHA), Home and Away shots (HAS), Home and Away corners (HAC), Home and Away Odds (HAOD), Home and Away attack strength (HAAT), Home and Away Players' performance index (HAPPI), Home and Away Managers' performance index (HAMPI), Home and Away streak (HASTK), Home and Away managers' win (HAMW). They used 20 matches

from the tenth and eleventh week of the 2014/15 English Premier League season to make their predictions. Using Logistic Regression and weighting to optimize characteristics, the outcome has an 85% accuracy.

Rotshtein et al. [13] developed a football fuzzy model that used a fusion of evolutionary pattern recognition and programs to adapt the parameters. They used this approach to analyze tournament data from Finland's national championship and discovered that these model parameter selections adjusting strategies. In summary, machine learning algorithms such as the Genetic Algorithm [14], Artificial Neural Networks [15], Support Vector Machine [16], K-Nearest Neighbors [17] or Bayesian Networks [18–20] have all been instrumental in predicting football match outcomes, attempting to account for a variety of factors that may influence the match results.

30.3 Experiment

This paper will compare the performance of function-based classification algorithms under the umbrella of Artificial Neural Networks, which are Multilayer Perceptron (MLP) and Dense Neural Networks (DNN). Both implementations of the algorithms are available under the Wekadeeplearning4j package in the Waikato Environment for Knowledge Analysis (WEKA) (see <https://deeplearning.cms.waikato.ac.nz/>). WEKA is an open source software under the General Public License (GNU), where this software may provide the implementation state-of-the-art data mining, machine learning and deep learning algorithm [21].

Apart from comparative experiments between the performance of ANN vs. DNN in football match classification, addition experiments were carried out to explore hyperparameter tuning in DNN. Different number of hidden layers were experimented along with different number of nodes (size of the layers) and different number of epochs.

30.3.1 Data Description

The dataset used in this study is sourced from a public website <http://www.football-data.co.uk/>. The dataset of English Premier League (EPL) as shown in the table covers the results from the previous ten seasons of EPL. The objective of the experiment is to classify the outcome of football matches based on the Full Time Result (FTR), where the values are Home Win (H), Away Win (A), and Draw (D). Table 30.1 shows some of the key factors and values used for football match prediction. In this case study, the data from season 2018–2019 was selected. This final selected dataset contains 62 attributes and 380 rows total of data.

Table 30.1 Sample dataset from the English Premier League

Label	Attribute	Sample values	Data type
HomeTeam	–	Man United	–
AwayTeam	–	Leicester	–
FTR	Full time result	H	Nominal
HS	Home team shots	8	Numeric
AS	Away team shots	13	Numeric
HST	Home team shots on target	6	Numeric
AST	Away team shots on target	4	Numeric
HF	Home team fouls committed	11	Numeric
AF	Away team fouls committed	8	Numeric
HC	Home team corners taken	2	Numeric
AC	Away team corners taken	5	Numeric
HY	Home team yellow cards	2	Numeric
AY	Away team yellow cards	1	Numeric
HR	Home team red cards	0	Numeric
AR	Away team red cards	0	Numeric

30.3.2 Classifier Models

Two classifier models used in the experiments are the Multilayer Perceptron (MLP) and the Dense Neural Networks (DNN). MLP is a Machine Learning algorithm that attempts to mimic the functionalities of artificial neural systems. It processes data using a connectionist approach to computing and is made up of a linked collection of basic processing components or artificial neurons. During the process of learning, MLP is thought to be an adaptive system that alters its structure in response to internally or externally information flowing through the networking.

In WEKA, the MLP classifier uses backpropagation to learn a multi-layer perceptron to classify instances. The learning rate and the momentum rate for the back-propagation algorithm is set to the default 0.3 and 0.2 respectively. The number of epochs is set to 100 for the comparative purposes. The number of hidden layers created for the network is set to ‘a’, which represents attributes and classes in WEKA. Figure 30.1 shows the excerpt of the network parameters.

The implementation of Dense Neural Networks (DNN) model used in this study is based on WekaDeeplearning4j deep learning package for the WEKA workbench. In this package, a DNN refers to one specific deep learning architecture, which is the DenseLayer. DenseLayer basically connects all units or nodes to its parent layer. The backend is provided by the Deeplearning4j Java library. DNN are layers in which all the neurons in a network layer are completely linked (dense). Each neuron in a layer receives input from every neuron in the layer before it. As a result, they’re intertwined. To put it simply, the dense layer is a completely linked layer, which means that all the neurons in one layer are linked to those in the next. The example of structure of the DNN is shown as Fig. 30.2.

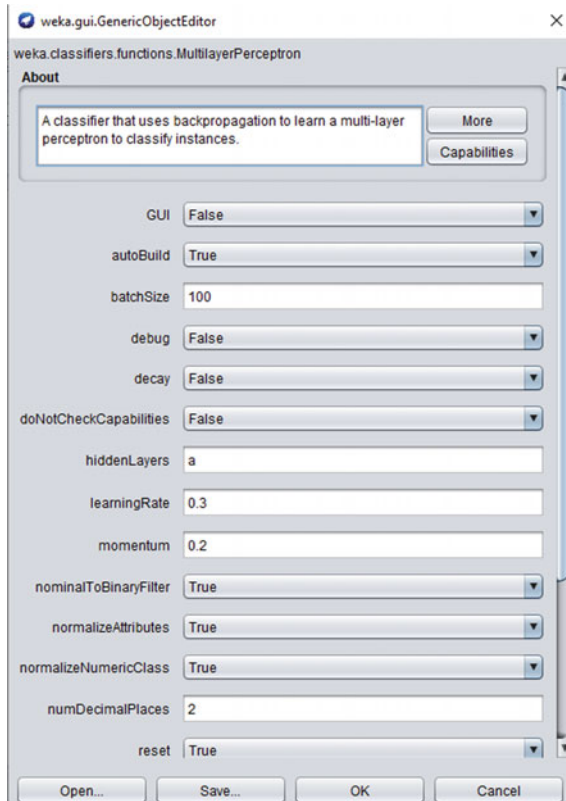


Fig. 30.1 MLP network configuration in WEKA

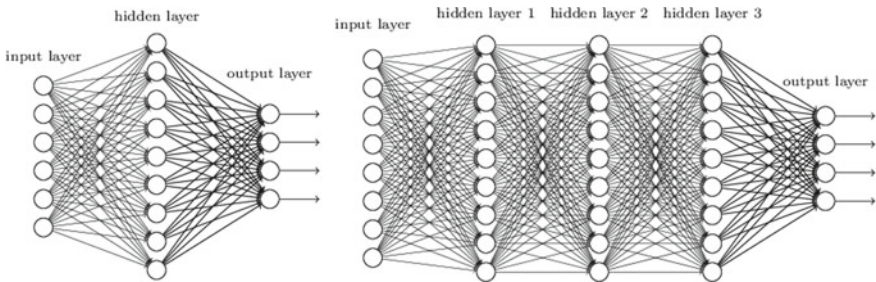


Fig. 30.2 The structure of ANN (non-deep feedforward neural networks) versus DNN (deep neural networks) [22]

The basic experimental setting for DNN will be using 8-node, single layer with 100 epochs run size. In the subsequent experiments, the layers are then extended up to two more layers, which are 8×8 -dense layer and $8 \times 8 \times 8$ -dense layer. Further, 16-node dense layers are investigated in order to see the impact of changing parameter in DNN. Note that the number of nodes or the size of intermediate layers follow geometric progression from 8 to 16.

30.3.3 Performance Evaluation

The performance of both MLP and DNN classifiers is the central focus of this paper. The measurement of Accuracy, Precision, Recall, and F-measure are used for comparison. WEKA produces a confusion matrix that will in turn provides the values for percentage rate for all 4 measurements. There are four possible results from this classification experiment from the confusion matrix, which are True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN). The first evaluation metric, Accuracy, is calculated as the total number of correct predictions divided by the total number of valid predictions in the dataset. The highest level of accuracy is 1.0, while the lowest level is 0.0. The formula is shown in Eq. 30.1.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{FP} + \text{TP} + \text{TN} + \text{FN}} \quad (30.1)$$

Next, Precision is calculated as the total number of positive prediction divided by the number of true positive forecasts It is also known as positive predictive value (PPV). The highest level of precision is 1.0, while the lowest level is 0.0. The formula is shown in Eq. 30.2.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (30.2)$$

Recall is calculated as the total number of positives divided by the number of true positive forecasts. It is also known as the true positive rate (TPR). The highest level of sensitivity is 1.0, while the lowest level is 0.0. The formula is shown in Eq. 30.3.

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} = \text{REC} = \frac{\text{TP}}{\text{P}} \quad (30.3)$$

Finally, F-measure is Accuracy and Recall at a harmonic mean. The formula is shown in Eq. 30.4.

$$\text{F-measure} = \frac{\text{TP}}{\text{TP} + \frac{1}{2}(\text{FP} + \text{FN})} \quad (30.4)$$

30.4 Results and Discussion

Table 30.2 shows the comparison of results between the both Multilayer Perceptron (MLP) and Dense Neural Networks (DNN) classifiers in the multi-class classification experiments of the football match English Premier League dataset. The performance evaluation for Home Win (H), Away Win (A), and Draw (D) classes are measured in terms of Accuracy, Precision, Recall, and F-Measure. In the experiment, 10-fold cross validation were used to conduct these experiments.

Figures 30.3 and 30.4 show the excerpt of experimental results from WEKA. From the table, we can see that the MLP classifier produced higher accuracy as compared to DNN with accuracy of 78.42%. The results for precision, recall, and F-measure are also closer to 1 as compared to DNN. DNN only achieved 67.64% accuracy.

To further investigate the performance of DNN, a subsequent comparative experiments were carried out to compare performance of DNN under different parameter settings. Because this is a multi-class classifier, the size of the output layer for both MLP and DNN classifiers are fixed to three, to represent H, A, and D. However, the experiments tested with different number of nodes in the intermediate layers. Meanwhile, the experiments also used two different epoch sizes; 100 and 200. The size of epoch is a hyperparameter that defines the number of times that both the MLP and DNN algorithms work through the entire training dataset. One epoch means training the neural network with all the training data for one cycle. Table 30.3 shows the complete results for variations of fully-connected network structure between one to three layers. Each fully connected layers are defined as a dense layers with either 8 or 16 nodes.

From Table 30.3, the 8-dense layers with 100 epochs achieved similar classification accuracy in first and third layers, which was 67.63%. The third layer, however, has a slightly higher precision and F-measure which was 0.678 and 0.677 respectively, as compared to the first layer precision and F-measure which was both 0.676. This indicates that the third layer for the 8-dense and 100 epochs of the DNN approach has the best accuracy performance for its category. Next, for the 8-dense and 200 epochs in DNN showed that the third layer has the highest accuracy among the first and second layer, which was 70.26%. Thus its precision, F-measure and Recall was also the highest which was 0.699, 0.701, 0.703 respectively. This indicates that the third layer for the 8 dense and 200 epochs of the DNN approach has the best accuracy for its category.

In addition, the category of 16-dense and 100 epochs of DNN, the results showed that the third layer has the highest accuracy, which was 70.26% compare to the other

Table 30.2 Comparative evaluation for classification performance evaluation of MLP versus DNN

Algorithms	Accuracy (%)	Precision	Recall	F-measure
MLP	78.42	0.780	0.784	0.782
DNN	67.63	0.676	0.676	0.676


```

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances      298          78.4211 %
Incorrectly Classified Instances    82          21.5789 %
Kappa statistic                    0.6531
Mean absolute error                 0.1555
Root mean squared error             0.3526
Relative absolute error             37.3208 %
Root relative squared error         77.2639 %
Total Number of Instances          380

=== Detailed Accuracy By Class ===

      TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
0.878  0.121  0.869    0.878  0.874    0.758  0.955    0.952    H
0.836  0.095  0.817    0.836  0.826    0.737  0.949    0.898    A
0.451  0.110  0.485    0.451  0.467    0.351  0.776    0.472    D
Weighted Avg.  0.784  0.110  0.780    0.784  0.782    0.674  0.919    0.844
    
```

Fig. 30.3 WEKA output for MLP algorithm

```

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances      257          67.6316 %
Incorrectly Classified Instances    123          32.3684 %
Kappa statistic                    0.4822
Mean absolute error                 0.2185
Root mean squared error             0.448
Relative absolute error             52.4268 %
Root relative squared error         98.1628 %
Total Number of Instances          380

=== Detailed Accuracy By Class ===

      TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
0.762  0.211  0.767    0.762  0.765    0.552  0.877    0.858    H
0.734  0.139  0.729    0.734  0.732    0.594  0.891    0.828    A
0.352  0.149  0.352    0.352  0.352    0.203  0.697    0.337    D
Weighted Avg.  0.676  0.175  0.676    0.676  0.676    0.501  0.848    0.751
    
```

Fig. 30.4 WEKA output for DNN algorithm

two layers, while its precision, F-measure and recall were 0.706, 0.704, and 0.703 respectively. This indicates that the third layer of the 16-dense and 100 epochs of the DNN approach has the highest accuracy as compared to the first and second layer. Finally, 16-dense layers and 200 epochs of DNN revealed a slightly different outcome as compared to the others. This is because the highest accuracy came from the second layer with an accuracy of 70.79%. However, this layer produced low precision and F-measure, which were both 0.706 as compared to the third layer precision and F-measure, which were 0.708 and 0.705 respectively. Since, there was only a slightly difference on the precision and F-measure, the second layer of the 16-dense and 200 epochs of the DNN approach has the highest accuracy as compared to the others approach.

Table 30.3 Performance evaluation for DNN classifier under different parameter settings

Dense layers	Epochs	Accuracy (%)	Precision	Recall	F-measure
8	100	67.63	0.676	0.676	0.676
8 × 8	100	64.74	0.661	0.647	0.654
8 × 8 × 8	100	67.63	0.678	0.676	0.677
8	200	68.16	0.689	0.682	0.685
8 × 8	200	64.47	0.662	0.645	0.652
8 × 8 × 8	200	70.26	0.699	0.703	0.701
16	100	67.89	0.674	0.679	0.676
16 × 16	100	70.00	0.703	0.700	0.701
16 × 16 × 16	100	70.26	0.706	0.703	0.704
16	200	66.84	0.660	0.668	0.664
16 × 16	200	70.79	0.706	0.708	0.706
16 × 16 × 16	200	70.53	0.708	0.705	0.707

Overall, we can see from the results that the accuracy rate slowly increasing as we add more layers to the DNN. The same effect can also be achieved if we increase the number of nodes or the size of the intermediate layers or the number of epochs in the experiments. For example, increasing the geometric progression from 8 to 16 dense layers produced the same accuracy rate of 70.26% with higher number of epochs (200 epochs for an 8-dense layers vs. 100 epochs for a 16-dense layers).

30.5 Conclusion

Match classification in football has become a topic of interest in recent years, particularly when it comes to match outcome prediction. Using data from the English Premier League, this research demonstrated how the Multilayer Perceptron (MLP) and the Dense Neural Networks (DNN) were used to classify the outcome of a football match. The experiment revealed that MLP produced better classification accuracy rate of 78.42% as compared to 67.63% by the standard DNN. Nonetheless, the performance of DNN can be further improved by means of hyperparameter tuning, whereby DNN achieved the highest accuracy of 70.53% when modeled in a three-dense layers with the size of 16 nodes and trained with 200 epochs. Although by increasing the number of nodes and size of epochs is able to increase the accuracy rate for football match results classification, this will greatly affect the computational time due to increased complexity of the network. In addition, the big number of epochs is not necessary unless the input data comes from a large dataset. The number of epochs is not significant as compared to the validation and training error. Training should stop as the validation error starts increasing, which means the

accuracy rate starts decreasing. This is a indication of overfitting. In conclusion, these results are hoped to serve as a baseline for future study into the classification of football match outcomes based on deep neural networks architecture.

Acknowledgements This research is supported by the Ministry of Higher Education (MOHE) through Fundamental Research Grant Scheme (FRGS/1/2020/ICT02/UTHM/02/4).

References

1. L.M. Hvattum, H. Arntzen, Using ELO ratings for match result prediction in association football. *Int. J. Forecast.* **26**(3), 460–470 (2010). <https://doi.org/10.1016/j.ijforecast.2009.10.002>
2. J. Goddard, I. Asimakopoulous, Forecasting football results and the efficiency of fixed-odds betting. *J Forecast.* **23**(1), 51–66 (2004). <https://doi.org/10.1002/for.877>
3. T.U. Grund, Network structure and team performance: the case of English premier league soccer teams. *Soc. Netw.* **34**(4), 682–690 (2012). <https://doi.org/10.1016/j.socnet.2012.08.004>
4. C. Lago-Peñas, A. Dellal, Ball possession strategies in elite soccer according to the evolution of the match-score: the influence of situational variables. *J. Hum. Kinet.* **25**(1), 93–100 (2010). <https://doi.org/10.2478/v10078-010-0036-z>
5. F.A. Moura, L.E.B. Martins, S.A. Cunha, Analysis of football game-related statistics using multivariate techniques. *J. Sports Sci.* **32**(20), 1881–1887 (2014). <https://doi.org/10.1080/02640414.2013.853130>
6. T. Kirschstein, S. Liebscher, Assessing the market values of soccer players-a robust analysis of data from German 1. and 2. Bundesliga. *J. Appl. Stat.* **46**(7):1336–1349 (2019). <https://doi.org/10.1080/02664763.2018.1540689>
7. A. Efrati, How “deep learning” works at apple, beyond (2021). <https://www.theinformation.com/articles/how-deep-learning-works-at-apple-beyond>
8. N. Jones, Computer science: the learning machines. *Nature* **505**(7482), 146–14, 843 (2014)
9. M.A. Rahman, A deep learning framework for football match prediction. *SN Appl Sci* **2**(2), 1–12 (2020)
10. M.C. Purucker, Neural network quarterbacking. *IEEE Potentials* **15**(3), 9–15 (1996). <https://doi.org/10.1109/45.535226>
11. J. Kahn, Neural network prediction of NFL football games. World Wide Web Electronic Publication, pp 9–15. <http://homepages.cae.wisc.edu/~ece539/project/f03/kahn.pdf>
12. I. Chinwe Peace, An improved prediction system for football a match result. *IOSR J. Eng.* **04**(12), 12–020 (2014). <https://doi.org/10.9790/3021-04124012020>
13. A.P. Rotshtein, M. Posner, A.B. Rakityanskaya, Football predictions based on a fuzzy model with genetic and neural tuning. *Cybern. Syst. Anal.* **41**(4), 619–630 (2005). <https://doi.org/10.1007/s10559-005-0098-4>
14. A.P. Rotshtein, M. Posner, A. Rakityanskaya, Football predictions based on a fuzzy model with genetic and neural tuning. *Cybern. Syst. Anal.* **41**(4), 619–630 (2005)
15. S. Guan, X. Wang, Optimization analysis of football match prediction model based on neural network. *Neural Comput. Appl.* 1–17 (2021)
16. C.P. Igiri, Support vector machine-based prediction system for a football match result. *IOSR J. Comput. Eng. (IOSR-JCE)* **17**(3), 21–26 (2015)
17. E. Esme, M.S. Kiran, Prediction of football match outcomes based on bookmaker odds by using k-nearest neighbor algorithm. *Int. J. Mach. Learn. Comput.* **8**(1), 26–32 (2018)
18. A.C. Constantinou, N.E. Fenton, M. Neil, Pi-football: a Bayesian network model for forecasting association football match outcomes. *Knowl.-Based Syst.* **36**, 322–339 (2012)
19. F. Owramipur, P. Eskandarian, F.S. Mozneb, Football result prediction with Bayesian network in Spanish League-Barcelona team. *Int. J. Comput. Theory Eng.* **5**(5), 812 (2013)

20. N. Razali, A. Mustapha, F.A. Yatim, R. Ab Aziz, Predicting football matches results using bayesian networks for English premier league (epl). in IOP Conference Series: materials Science and Engineering, no. 226. (IOP Publishing, 2017), p. 012099
21. S. Lang, F. Bravo-Marquez, C. Beckham, M. Hall, E. Frank, Wekadeeplearning4j: a deep learning package for weka based on deeplearning4j. Knowl.-Based Syst. **178**, 48–50 (2019). <https://doi.org/10.1016/j.knosys.2019.04.013>. <http://www.sciencedirect.com/science/article/pii/S0950705119301789>
22. M. Nielsen, Neural networks and deep learning (2021). <http://neuralnetworksanddeeplearning.com/index.html>

Chapter 31

Comparative Analysis of Statistical and Machine Learning Methods for Classification of Match Outcomes in Association Football



Syazira Zulkifli, Aida Binti Mustapha, Shuhaida Ismail, and Nazim Razali

Abstract Association football is one of the most popular sports in the world, having a large fan base that draws media and entertainment platforms to it. The three major difficulties to developing highly accurate match result prediction models, especially in prediction match outcomes, are data availability and quality [1], model assumptions [2], and testing various models and parameters [3]. The primary goal of the study is to identify the best model in predicting football match outcomes. The football dataset was obtained from the top 5 leagues in Euro. Exploratory data analysis had been conducted to better understand the dataset. Models used in predicting the football match outcomes include Logistic Regression, Artificial Neural Networks, and XGBoost. The predictive performance of three classification models was compared in terms of accuracy, precision and recall. The results showed that the Artificial Neural Networks achieved highest accuracy of 0.6788, followed by Logistic Regression (0.668) and XGBoost (0.654). These results are hoped to be used as benchmark results for future experiments in the area of football match classification.

31.1 Introduction

Association football is one of the most popular sports, having a large fan base that draws media and entertainment attention. Match forecasts may give significant information to commentators and journalists, as well as vital inputs to sport decision-makers such as the team management, coaches or even spectators for betting. Therefore, in-game prediction is beneficial for both entertainment and risk assessment for

S. Zulkifli · A. B. Mustapha (✉) · S. Ismail
Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia,
84600 Johor, Panchor, Malaysia
e-mail: aidam@uthm.edu.my

N. Razali
Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn Malaysia,
86400 Batu Pahat, Johor, Parit Raja, Malaysia

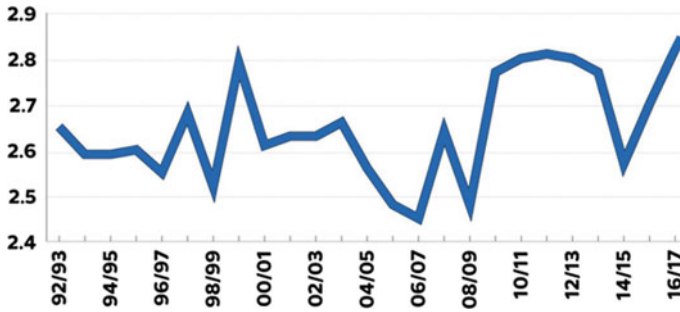


Fig. 31.1 Average number of goals scored per premier league game

players during the match. In addition, results of sports events can be difficult to predict, and miracles often occur. The ball in particular is an interesting example due to its strict match length as opposed to racket games like tennis, where the game is played until the player wins. It also has one type of scoring event, which is the goals. This is different from a rugby match where different score carries different points, hence there is an unlimited number of scoring event times in the middle of the game, and all worth the same amount.

A team's performance in a football match might result in a victory, a defeat, or a draw. As a result, predicting the outcome of a game may seem to be rather simple. Traditional predictive approaches have relied on match results to assess team performance and create statistical models to forecast future game outcomes. However, the average number of goals scored per game in the English Premier League during the last 15 years is rather low (less than 3 goals per game on average) due to the low-scoring character of games (<https://footystats.org/england/premier-league/average-total-goals-table>). As shown in Fig. 31.1, the number of goals scored during a match has a random element. For example, a team with several scoring chances may be unfortunate and fail to convert any of them into goals, but a side with a single scoring chance may score. As a consequence, match results are an imprecise assessment [4] of a team's performance and hence an incomplete tool for predicting future outcomes.

The use of in-game data to go beyond the plain match outcomes may be considered as a possible solution to this issue. Examples of in-game data collected during a football games include player performance, the quality of shots, on the ball actions, dribbles and set pieces. Because in-depth match data have been made public in recent years, researchers have started to explore beyond the match outcome [5]. As a result, predicted goals measures have been developed, where estimate the amount of goals a side is calculated to achieve in a game in order to eliminate the random aspect of goalscoring.

Along with the development in game prediction in sports analytics, Machine Learning (ML) approaches have also rise to support this trend by improving prediction performance in a variety of classification [6] and regression issues [7]. The

development of better models for both predicting the result of a match and predicting the actual score have been facilitated by the study of these distinct methodologies and algorithms such as Bayesian networks [8], Poisson autoregression with exogenous covariates PARX model [9], and ML [10].

In sports analytics, particularly in predicting match outcomes, there are three main challenges to building a highly accurate match outcome prediction models, which are data availability and quality [1], model assumptions [2] and testing different models and parameters [3].

- **Data availability and quality:** Prediction tasks in match outcomes requires deep mathematical analysis but public databases mostly provide basic data such as match statistics and individual player rating. Various details of public football matches must be assessed in order to find the right one for us to use. Lack of centralized database is also another main issue [11], hence web scraping techniques to source for the dataset is necessary.
- **Research and understanding of the world of speculation:** In order to build our models and test different assumptions, detailed background research on predictive strategies and improve mathematical understanding of the various machine learning algorithms are required before the techniques can be used in our predictions.
- **Testing different models and parameters:** A comprehensive machine learning pipeline is important to test different models and parameters. This way testing can be performed in shorter amount of time in effort to obtain the best models.

In building a relatively high accuracy of match outcome prediction model, this paper is set to explore football data to be used as features in the match outcome prediction model and apply three classification techniques for the match outcome prediction model, which are Logistic Regression (LR), Artificial Neural Network (ANN), and Extreme Gradient Boosting (XGBoost) Algorithm.

The remaining of this paper proceeds as follows. Section 31.2 presents the related works, Sect. 31.3 presents the research methodology, Sect. 31.4 presents the evaluation results, and finally Sect. 31.5 concludes the paper.

31.2 Related Work

For the prediction of football outcomes [12], utilized a Bayesian hierarchical model. In recent years, the topic of football modelling has grown in popularity, and numerous different models have been developed to predict the characteristics that cause a team to lose or win a game, as well as to anticipate a game's outcome. To achieve these two goals and to test the Bayesian hierarchical model, data from the 1991–1992 Italian Series A championship was employed. To tackle the over-reduction issue caused by the Bayesian hierarchical model, they offer a more sophisticated mix model that fits the actual data. The 2007–2008 Italian Series A championship served as the basis for benchmark comparisons.

Machine learning approaches were employed by [13] to forecast football results. The main purpose of this study is to put a variety of machine learning approaches to the test in order to predict the end and result of football matches based on in-game match activities rather than the number of goals scored by each team. They've put several model architectural ideas to the test and compared the efficiency of their models to benchmarking tools. Rather than using specific realized objectives, they developed an "expected objective" metric in this study to help us analyze team performance. To develop a classification model that predicts the results of future matches and a regression model that predicts future matches, this measure is combined with the measurement of an offensive and defensive team ranking update throughout the game. Their models' efficiency is comparable to that of bookies and corresponds well with existing popular techniques.

Tax and Joustra [14] developed a machine learning strategy for forecasting the Dutch football tournament using public data. This article defines a framework for the Dutch Eredivisie that focuses on public data. The determinants of predictive usefulness for match outcomes were outlined in a comprehensive literature review. Characteristics of the candidates have been developed. Modeling preparation was supplemented by self-made public data collection, which consisted of 13 Dutch Eredivisie match data seasons. On public data, a range of modifications in dimensional reduction methods and classification algorithms have been examined. The greatest detection precision for the public data feature collection was achieved using a combination of PCA (with a 15% difference) and a Naive Bayes or Multilayer Perceptron classifier. Models for betting odds have been developed, as well as a hybrid feature set (common data union and wagering odds characteristics).

Baboota and Kaur [15] used a machine learning methodology to forecast and model football outcomes for the English Premier League. This research shows their efforts to create a standard statistical model for English Premier League games. They created a feature collection utilizing software engineering and an exploratory data analysis, which assesses the key elements for forecasting football match results and uses machine learning to create a highly comprehensive prediction framework. They showed that the success of their concept is predicated on a number of key criteria. Their best model with gradient boosts achieved a performance of 0.2156 in the probability (RPS) metric for game weeks 6 to 38 in the EPL aggregated over two seasons (2014–2015 and 2015–2016), whereas the betting organizations they consider (Bet365 and Pinnacle Sports) received RPS values of 0.2012 for the same period. Because the low RPS value indicates a greater predictive accuracy, despite the positive results, their model did not outperform the bookmaker's predictions.

Predicting Football Match Results using a Logistic Regression Model was created by [16]. They designed it to anticipate match outcomes in the Barclays Premier League season 2015/2016 for home and away wins, as well as to figure out what the key factors are in winning matches. Their work differs from others in that they only employ significant factors derived from study in the same subject, rather than speculating on what the relevant factors may be. They also included data from the video game FIFA, since [17] shown that including data from the game may increase prediction quality. The model was created utilising training data from the 2010/2011

season through the 2015/2016 season. Logistic regression is a classification approach that may be used to predict sports outcomes and provide extra information through regression coefficients. “Home Offense”, “Home Defense”, “Away Offense”, and “Away Defense” are the variables employed. They experimented by changing the seasons in which the training data was collected. The constructed model has a prediction accuracy of 69.5 percent.

Rahman [18] used publicly accessible datasets, which are the International Football outcomes from 1872 to 2018, FIFA Soccer Ranking, FIFA World Cup 2018 Dataset, and FIFA Soccer Ranking datasets to estimate the outcome of the FIFA World Cup 2018 matches in 2020. The authors suggested an ANN and LSTM-based model that could accurately predict the result of the matches by 63.3%. LSTM was utilized by [19] to predict match results. The writers compiled a dataset from a variety of sources and classified it based on player statistics and match history. Matches from several European Leagues were included in the dataset, which spanned the years 2011–2016 at 52.4% accuracy rate.

Blaikie et al. [20] utilized ANN to forecast football results in the NFL and NCAA. They looked through the data to find the best predictive statistics to utilize in the model. The NFL model outperformed the NCAA model when derivative analysis was used. When compared to other prediction experts, the NFL model regularly placed in the top half. Over the course of five seasons [21], employed ANN and other prediction algorithms to forecast the score result of matches played in the Spanish La Liga. They projected the result of home and away matches using the FIFA 18 game database, which included match history and Team against Team databases, as well as individual career information. With LR, they were able to obtain an accuracy of 71.63%. Their ANN model has a match history database accuracy of 63.1% and a combination of the match history database and the Team against Team database accuracy of 69.2%. Artificial Neural Networks were used by [22] to forecast team standings in the male professional volleyball league. The data included in the research was collected during a 2-year period (2013–2015). Using a layer 4-neuron model with the “logsig” transfer function, “trainlm” training function, and “learned” adaptive learning function, the accuracy was 98%.

Chen and Guestrin [23] were the first to develop the XGBoost method, which improved on the basic gradient boosted decision trees by upgrading the method and optimizing the system. This approach was created to increase model performance and execution speed. On classification and regression predictions, the XGBoost method performs well with structured or tabular information. This method also has a regularized goal for greater generalization and an additive solution for the generic objective (cost) functions [24]. Using the XGBoost Algorithm [25], developed a machine learning strategy for football prediction. They provided a model for prediction in this research that used the XGBoost algorithm as the core method. The method is compared to a variety of different algorithms to evaluate how well it performs. The research also claimed that the suggested model could outperform several other models using just public data.

31.3 Methodology

This paper proposes a comparative experiments for predicting match outcomes in Association Football. Figure 31.2 shows the research framework for the implementation of the experiment.

From the figure, modeling will be performed using three models, which are Artificial Neural Networks (ANN), XGBoost, and Logistic Regression (LR) coded using Python libraries in the Jupyter Noteboook. All experiments used the standard parameters in each Phytion libraries; ANN, XGBoost and LR. The performance of all models will be compared based on accuracy, precision, recall, and F-measure.

31.3.1 Dataset

The dataset was obtained from a public United Kingdom based source, from <http://www.football-data.co.uk/website>. This dataset contains data for the last 10 seasons of the German Bundesliga (<https://datahub.io/sports-data/german-bundesliga>) including the current season. The used dataset includes all five leagues but is limited to eight features: home team, away team, full time home team goals (FTHG), full time away team goals (FTAG), full time result (FTR), half time home team goals (HTHG), half

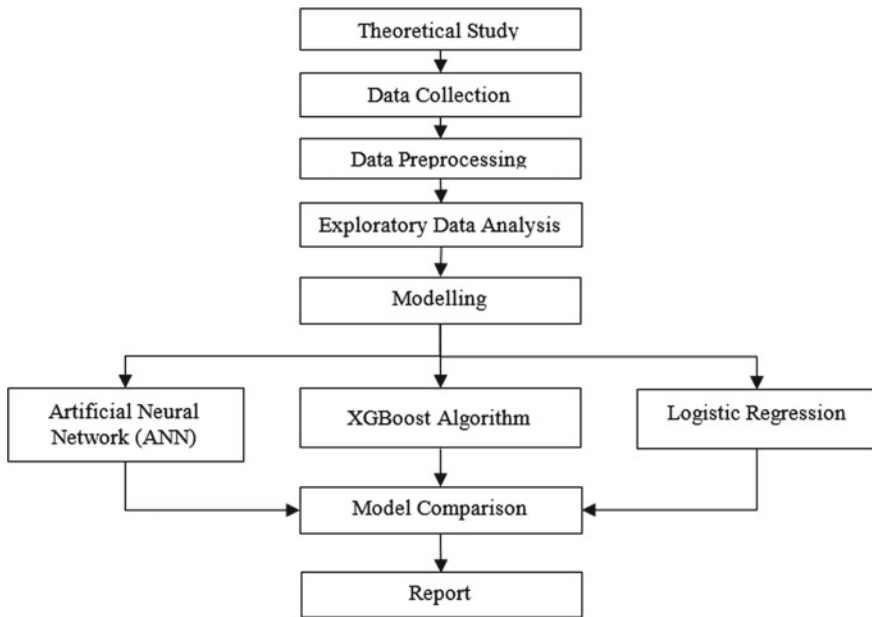


Fig. 31.2 Research framework

Div	Date	HomeTeam	AwayTeam	FTHG	FTAG	FTR	HTHG	HTAG	HTR	Referee	HS	AS	HST	AST	HF	AF	HC	AC	HY	AY
E0	31/12/16	Southampton	West Brom	1	2	A	1.0	1.0	D	M Jones	10.0	7.0	1.0	3.0	13.0	17.0	8.0	3.0	1.0	3.0
E0	31/12/16	Chelsea	Stoke	4	2	H	1.0	0.0	H	R Madley	18.0	5.0	7.0	2.0	12.0	10.0	8.0	3.0	3.0	2.0
E0	31/12/16	Swansea	Bournemouth	0	3	A	0.0	2.0	A	K Friend	14.0	11.0	5.0	6.0	14.0	9.0	4.0	3.0	2.0	0.0
E0	31/12/17	Crystal Palace	Man City	0	0	D	0.0	0.0	D	J Moss	10.0	15.0	1.0	4.0	14.0	11.0	7.0	8.0	5.0	4.0
E0	31/12/17	West Brom	Arsenal	1	1	D	0.0	0.0	D	M Dean	14.0	14.0	3.0	4.0	14.0	9.0	5.0	6.0	3.0	3.0

Fig. 31.3 Sample of dataset

time away team goals (HTAG), and half time result (HTR). Excerpt of the dataset is shown in Fig. 31.3.

Data preprocessing is a stage in the machine learning process in which data is cleaned, filtered, scaled, or encoded so that the algorithm can better understand the data’s features. In this study, the predictions for football match outcomes dataset obtained will be cleaned, encoded and standardized before fitting them into different models. After preprocessing the predictions for football match outcomes dataset, exploratory data analysis will be conducted to visualize and gain insights of the data as well as to further comprehend the data’s distribution and characteristics.

The top five leagues are represented in the dataset. Using statistics from the top five famous leagues during the past ten years, a total of 38 matches are played by 20 teams each, for a total of 380 teams in a single season. As a result, there are a total of 17,880 matches. Looking at certain datasets, it is obvious that home teams have an edge in winning since they have a larger issue count. Categorical variables are divided into three categories: ‘H’, ‘A’, and ‘D’. Based on Fig. 31.4, it is shown that the counts are not equal, hence this dataset is imbalanced.

31.3.2 Statistical and Machine Learning Algorithms

The algorithms used in the experiments come from two categories; statistical (Logistic Regression) and machine learning (Artificial Neural Networks and XGBoost). Implementations of the algorithms are sourced from Jupyter Notebook using all standard parameter settings. The following subsections present the theoretical brief of each algorithm.

Logistic Regression (LR) Algorithm The Logistic Regression (LR) function is shown in Fig. 31.5.

In LR, the weighted sum of inputs is processed through an activation function that may translate values between 0 and 1. The sigmoid function is a kind of activation function, and the curve it produces is known as the sigmoid curve or S-curve. The equation for logistic regression is shown in Eq. 31.1.

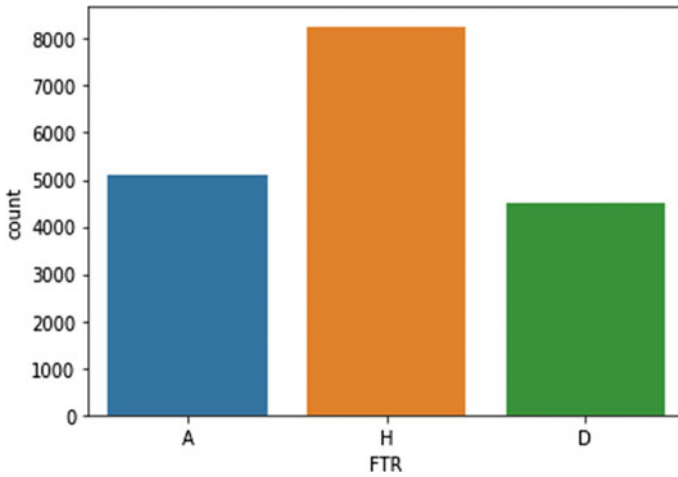
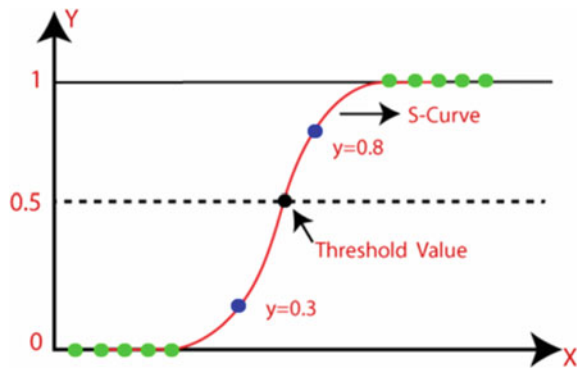


Fig. 31.4 Full-time result count

Fig. 31.5 Logistic regression function

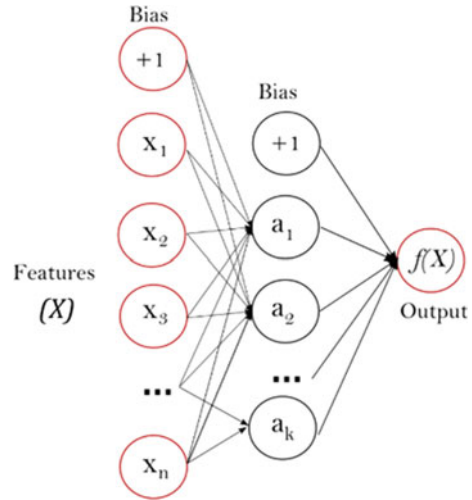


$$\log\left[\frac{y}{1-y}\right] = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + \dots + b_nx_n \quad (31.1)$$

In summary, any classification model for a binary output variable inherently uses logistic regression. In fact, the model’s output variable is the log-odds score of the classification’s good result. With the aid of independent factors, Logistic Regression (LR) is utilized to predict the categorical dependent variable. Only 0 and 1 may be the outcome of a Logistic Regression issue. The notion of Maximum Likelihood estimation underpins logistic regression.

Artificial Neural Network (ANN) Algorithm Artificial neural networks (ANNs) are a computer methodology based on how the brain solves issues. The human brain is made up of neurons, which are nerve cells that are linked together by axons. ANNs are made up of numerous nodes that resemble actual neurons in the brain. The neurons are linked together by linkages that resemble actual axons, and they

Fig. 31.6 Neural networks diagram



interact. Each node receives data from other nodes, conducts a basic operation, and then transmits the result to the next node. An ANN self-learns and so excels in areas where a typical programming technique fails to describe the answer. The design of a neural network with one hidden layer is shown in Fig. 31.6.

Based on the figure, x_1, x_2, \dots, x_n are input variables and a_j is the activation of unit k in layer j . The relationships between neuron x to neuron a carries the weights (w_{nk}) of respective inputs. The bias is a sum between the weighted inputs to form the net inputs. Bias and weights are both adjustable parameters of the neuron but the output of a neuron can range from negative infinity to positive infinity. Therefore a mechanism to map the inputs (X) to outputs $f(X)$ are required, which is through an activation function.

Extreme Gradient Boosting (XGBoost) Algorithm Extreme Gradient Boosting is abbreviated as XGBoost. There are minimal frills in the library since it is laser focused on computing speed and model performance. It does, however, provide a lot of complex methods. As it was designed and developed for the primary goal of model performance and computational speed, XGBoost has proved to push the boundaries of processing power for boosted trees algorithms. It was specifically designed to make the most of every byte of memory and hardware resource available for tree boosting methods.

31.4 Performance Evaluation

The performance of LR, ANN, and XGBoost algorithm is the focus of this research paper. The measurement of Accuracy (acc), Precision (prec), Recall (rec) and F-measure are used to analyses the prediction's accuracy. The following equations are

used to calculate the estimation. Accuracy (acc) is calculated as the total number of correct predictions divided by the total number of valid predictions in the dataset. The highest level of accuracy is 1.0, while the lowest level is 0.0. The formula is shown in Eq. 31.2 where TP is the true positive rate, TN is the true negative rate, FP is the false positive rate, and FN is the false negative rate.

$$\text{acc} = \frac{\text{TP} + \text{TN}}{\text{FP} + \text{TP} + \text{TN} + \text{FN}} = \frac{\text{TP} + \text{TN}}{\text{P} + \text{N}} \quad (31.2)$$

Precision (prec) is calculated as the total number of positive forecasts divided by the number of true positive forecasts. It's also known as positive predictive value (PPV). The highest level of precision is 1.0, while the lowest level is 0.0. The formula is shown in Eq. 31.3.

$$\text{prec} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (31.3)$$

Recall (rec) is calculated as the total number of positives divided by the number of true positive forecasts. It's also sensitivity is 1.0, while the lowest level is 0.0. The formula is shown in Eq. 31.4.

$$\text{rec} = \frac{\text{TP}}{\text{TP} + \text{FN}} = \frac{\text{TP}}{\text{P}} \quad (31.4)$$

F-measure is accuracy and recall at a harmonic mean. The formula is shown in Eq. 31.5. The perfect F1 score is 1 when both recall and precision are also perfect (score is 1). If the precision or recall is 0, F1 score would be 0.5. This means F1 score provides a single value that weights both precision and recall in a balanced way especially when working with imbalanced datasets.

$$\text{F-measure} = \frac{2 \times \text{prec} \times \text{rec}}{(\text{prec} + \text{rec})} \quad (31.5)$$

31.5 Result and Discussion

This sections report the results for Logistic Regression, Artificial Neural Networks, XGBoost algorithm and GridSearchCV for the top five leagues of football dataset in Europe. From Fig. 31.7, the result shows the experiment had been conducted to evaluate the accuracy performance for Logistic Regression.

From the results, it is showed that the accuracy result for Logistic Regression is 0.668 or 67% with 60% and 40% split of training and testing the model, respectively. It has the same value for weighted average precision and F1-measure, which is 0.60 (60%) while the value of recall is 0.67 (67%). It can be seen above precision and recall are good for Home team and Away team wins. Hence, it is good precision and

```

model = LogisticRegression(multi_class='ovr', max_iter=500)
model.fit(X_train, Y_train)

print("Train model score: %.3f" % model.score(X_train, Y_train))
print("Validation model score: %.3f" % model.score(X_test, Y_test))
y_pred = model.predict(X_test)
print(classification_report(Y_test, y_pred))

```

	precision	recall	f1-score	support
A	0.66	0.76	0.70	998
D	0.45	0.16	0.24	869
H	0.71	0.87	0.78	1709
accuracy			0.67	3576
macro avg	0.60	0.60	0.57	3576
weighted avg	0.63	0.67	0.63	3576

Fig. 31.7 Results for logistic regression

```

Epoch 1/10
287/287 [=====] - 2s 4ms/step - loss: 0.8266 - accuracy: 0.6134 -
val_loss: 0.7505 - val_accuracy: 0.6516
Epoch 2/10
287/287 [=====] - 1s 3ms/step - loss: 0.7332 - accuracy: 0.6649 -
val_loss: 0.7515 - val_accuracy: 0.6546
Epoch 3/10
287/287 [=====] - 1s 3ms/step - loss: 0.7282 - accuracy: 0.6656 -
val_loss: 0.7464 - val_accuracy: 0.6549
Epoch 4/10
287/287 [=====] - 1s 3ms/step - loss: 0.7241 - accuracy: 0.6690 -
val_loss: 0.7438 - val_accuracy: 0.6619
Epoch 5/10
287/287 [=====] - 1s 3ms/step - loss: 0.7235 - accuracy: 0.6684 -
val_loss: 0.7492 - val_accuracy: 0.6535
Epoch 6/10
287/287 [=====] - 1s 3ms/step - loss: 0.7077 - accuracy: 0.6737 -
val_loss: 0.7501 - val_accuracy: 0.6530
Epoch 7/10
287/287 [=====] - 1s 3ms/step - loss: 0.7124 - accuracy: 0.6735 -
val_loss: 0.7509 - val_accuracy: 0.6558
Epoch 8/10
287/287 [=====] - 1s 3ms/step - loss: 0.7092 - accuracy: 0.6670 -
val_loss: 0.7496 - val_accuracy: 0.6555
Epoch 9/10
287/287 [=====] - 1s 3ms/step - loss: 0.7167 - accuracy: 0.6707 -
val_loss: 0.7527 - val_accuracy: 0.6541
Epoch 10/10
287/287 [=====] - 1s 3ms/step - loss: 0.7113 - accuracy: 0.6788 -
val_loss: 0.7558 - val_accuracy: 0.6544
<keras.callbacks.History at 0x7f02ae082e50>

```

Fig. 31.8 Results for artificial neural networks

recall but for Draw the model has low F1-measure. Next, Fig. 31.8 shows the result for Artificial Neural Networks (ANN). The neural network was run for 10 epochs i.e. 10 iterations on the dataset used. It can be seen that the loss is getting reduced a bit and the accuracy of the model is close to previous model had, which is 0.6788 (68%). It is found that ANN does not provide much of an improvement as compared to the previous models.

```

from xgboost import XGBClassifier
model = XGBClassifier(n_jobs=-1)

model.fit(X_train, Y_train);

print("Train model score: %.3f" % model.score(X_train, Y_train))
print("Validation model score: %.3f" % model.score(X_test, Y_test))
y_pred = model.predict(X_test)
print(classification_report(Y_test, y_pred))

```

Train model score: 0.676					
Validation model score: 0.654					
	precision	recall	f1-score	support	
A	0.65	0.71	0.68	1012	
D	0.48	0.21	0.29	923	
H	0.69	0.87	0.77	1641	
accuracy			0.65	3576	
macro avg	0.61	0.60	0.58	3576	
weighted avg	0.62	0.65	0.62	3576	

Fig. 31.9 Result for XGBoost algorithm

Finally, the experiment was performed to assess the accuracy performance for XGBoost algorithm, as shown in Fig. 31.9. Based on the figure, the accuracy result for XGBoost algorithm is 0.654 (65%) with 60% and 40% split of training and testing the model respectively for experimental setting. The weighted average for Precision and F1-measure have the same value of 0.62 (65%), whereas recall has a value of 0.65 (65%). A special case of boosting where errors are minimized by gradient descent algorithm e.g. the strategy consulting firms leverage by using case interviews to weed out less qualified candidates.

Figures 31.10 and 31.11 show the GridSearchCV on the XGBoost model for testing. It is the process of performing hyperparameter tuning in order to determine the optimal values for a given model. As mentioned above, the performance of a model significantly depends on the value of hyperparameters. Note that there is no way to know in advance the best values for hyperparameters so ideally, all possible values need to test to know the optimal values. The improvement as all the results in cross validation are not more than 66% which is same as the default results but still a good method to perform to understand the techniques of Machine Learning.

31.6 Conclusions and Recommendations

This study is motivated by the work of [10] in predicting match outcomes in association football. This study aim to contribute to this area of research by exploring football data that used as features in the match outcome prediction model. Three classification techniques were applied for the match outcome prediction model which are

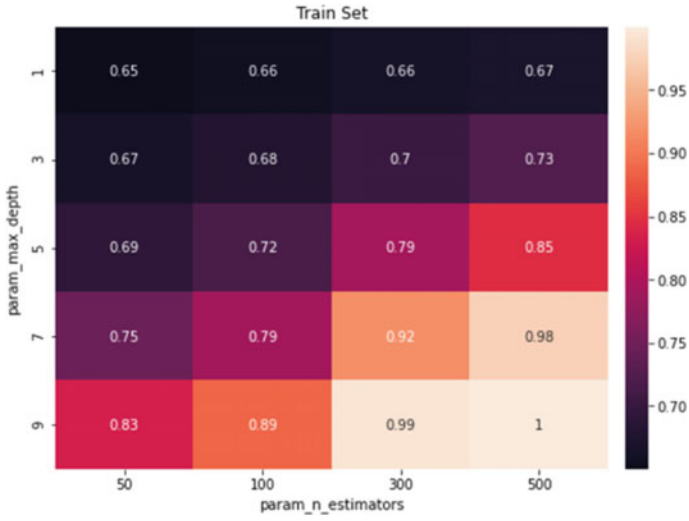


Fig. 31.10 Results for training set

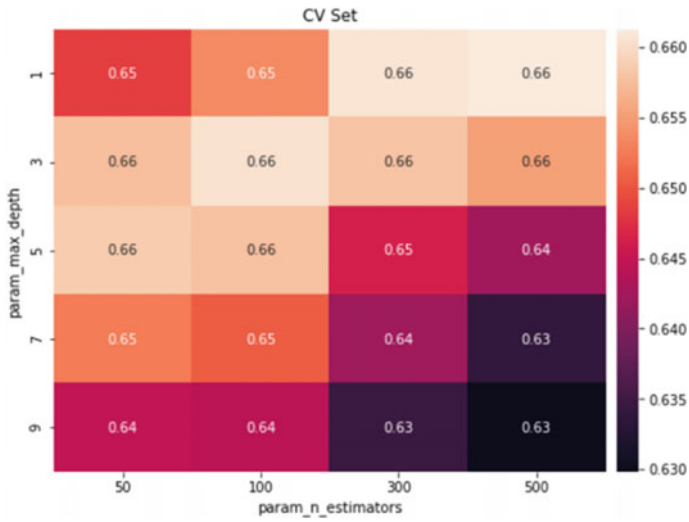


Fig. 31.11 Results for testing set

Logistic Regression (LR), Artificial Neural Network (ANN), and Extreme Gradient Boosting (XGBoost) Algorithm using the features formulated.

Another purpose for this study was to evaluate and compare the performance for all models developed based on their prediction accuracy, precision, and recall. Besides, data pre-processing has been carried out and relevant features were selected for the objective. The experiments showed that the Artificial Neural Networks achieved highest accuracy of 0.6788 (68%), followed by Logistic Regression (0.668) and XGBoost (0.654). To enhance additional accuracy, the precision and recall of draws must be improved. Other than that, technique as cross validation and hyper-parameter tuning were used to improve the accuracy of the models.

For the recommendations, this study may be used to build future work by exploring and fine-tuning the parameters of XGBoost algorithm so that the effect of different parameter values can be compared and it might potentially reduce the training time of the model too. This can further improve the effectiveness for real-time predictions. To further gain insights on only the significant variables contributing significantly to the models, feature importance selection can also be applied in future work. Besides, further work may consider integrating fuzzy logic or a hybrid to increase the diversity of the tress and flexibility of imperfect data management. The model can be considered to be embedded into an online portal that could automate and generate predictive analytic for future project.

Acknowledgements This research is supported by the Ministry of Higher Education (MOHE) through Fundamental Research Grant Scheme (FRGS/1/2020/ICT02/UTHM/02/4).

References

1. E. Morgulev, O.H. Azar, R. Lidor, Sports analytics and the big-data era. *Int. J. Data Sci. Anal.* **5**(4), 213–222 (2018). <https://doi.org/10.1007/s41060-017-0093-7>
2. A.C. Constantinou, Dolores: a model that predicts football match outcomes from all over the world. *Mach. Learn.* **108**(1), 49–75 (2019)
3. H. Arntzen, L.M. Hvattum, Predicting match outcomes in association football using team ratings and player ratings. *Stat. Modell.* **2020**, 1–22 (2020). <https://doi.org/10.1177/1471082X20929881>
4. M. Ahin, R. Erol, Prediction of attendance demand in European football games: comparison of ANFIS, Fuzzy Logic, and ANN. *Comput. Intell. Neurosci.* **2018** (2018). <https://doi.org/10.1155/2018/5714872>
5. L. Bransen, J. Van Haaren, Measuring football players' on-the-ball contributions from passes during games, vol. 11330 LNAI. Springer International Publishing (2019). https://doi.org/10.1007/978-3-030-17274-9_1
6. K. Kapadia, H. Abdel-Jaber, F. Thabtah, W. Hadi, Sport analytics for cricket game results using machine learning: an experimental study. *Appl. Comput. Inform.* (xxxx) (2019). <https://doi.org/10.1016/j.aci.2019.11.006>
7. F. Thabtah, L. Zhang, N. Abdelhamid, NBA game result prediction using feature analysis and machine learning. *Ann. Data Sci.* **6**(1), 103–116 (2019). <https://doi.org/10.1007/s40745-018-00189-x>

8. N. Razali, A. Mustapha, F.A. Yatim, R. Ab Aziz, Predicting football matches results using bayesian networks for english premier league (EPL). *IOP Conf. Ser.: Mater. Sci. Eng.* **226**(1) (2017). <https://doi.org/10.1088/1757-899X/226/1/012099>
9. G. Angelini, L. De Angelis, PARX model for football match predictions. *J. Forecast.* **36**(7), 795–807 (2017). <https://doi.org/10.1002/for.2471>
10. C. Herbinet, Predicting football results using machine learning techniques, in *2011 Proceedings of the 34th International Convention MIPRO*, vol. 48, pp. 1623–1627 (2018)
11. T. Blobel, M. Lames, A concept for club information systems (CIS) an example for applied sports informatics. *Int. J. Comput. Sci. Sport* **19**(1), 102–122 (2020). <https://doi.org/10.2478/ijcss-2020-0006>
12. G. Baio, M. Blangiardo, Bayesian hierarchical model for the prediction of football results. *J. Appl. Stat.* **37**(2), 253–264 (2010). <https://doi.org/10.1080/02664760802684177>
13. J. Hucaljuk, A. Rakipović, Predicting football scores using machine learning techniques, in *MIPRO 2011—34th International Convention on Information and Communication Technology, Electronics and Microelectronics—Proceedings*, vol. 48, pp. 1623–1627 (2011)
14. N. Tax, Y. Joustra, Predicting the Dutch football competition using public data: a machine learning approach. *Trans. Knowl. Data Eng.* **10**(10), 1–13 (2015)
15. R. Baboota, H. Kaur, Predictive analysis and modelling football results using machine learning approach for English Premier League. *Int. J. Forecast.* **35**(2), 741–755 (2019). <https://doi.org/10.1016/j.ijforecast.2018.01.003>
16. D. Prasetyo, Harlili: predicting football match results with logistic regression, in *4th IGNITE Conference and 2016 International Conference on Advanced Informatics: concepts, Theory and Application, ICAICTA 2016*, pp. 2–6 (2016). <https://doi.org/10.1109/ICAICTA.2016.7803111>
17. R. Gasparyan, *A Novel Way to Soccer Match Prediction*. Stanford University, Department of Computer Science (2014)
18. M.A. Rahman, A deep learning framework for football match prediction. *SN Appl. Sci.* **2**(2) (2020). <https://doi.org/10.1007/s42452-019-1821-5>
19. N. Danisik, P. Lacko, M. Farkas, Football match prediction using players attributes, in *DISA 2018—IEEE World Symposium on Digital Intelligence for Systems and Machines, Proceedings*, pp. 201–206 (2018). <https://doi.org/10.1109/DISA.2018.8490613>
20. A.D. Blaikie, G.J. Abud, J.A. David, R.D. Pasteur, NFL & NCAA football prediction using artificial neural networks, in *Proceedings of the Midstates Conference for Undergraduate Research in Computer Science and Mathematics*. Denison University, Granville, OH (2011). <http://ohio5.openrepository.com/ohio5/handle/2374.DEN/3930>
21. N. Zaveri, U. Shah, S. Tiwari, P. Shinde, L.K. Teli, Prediction of football match score and decision making process. *Int. J. Recent Innov. Trends Comput. Commun.* **6**(2), 162–165 (2018)
22. A.E. Tümer, S. Koçer, Prediction of team league’s rankings in volleyball by artificial neural network method. *Int. J. Perform. Anal. Sport* **17**(3), 202–211 (2017). <https://doi.org/10.1080/24748668.2017.1331570>
23. T. Chen, C. Guestrin, XGBoost: a scalable tree boosting system, in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2016)* (2016)
24. J. Brownlee, A gentle introduction to xgboost for applied machine learning. *Mach. Learn. Mastery* (2016). <http://machinelearningmastery.com/gentle-introduction-xgboost-appliedmachine-learning/>. Accessed 2 Mar. 2018
25. W. Gourh, K. Poojary, M. Vengarai, N. Parkar, Football prediction using XGBoost algorithm: a literature review. *J. Phys. Sci. Technol.* **12**(1), 109–112 (2020)

Part IV
Engineering

Chapter 32

Pre-assessment of Machine Learning Approaches for Patient Length of Stay Prediction



Farhanahani Mahmud , Ahmad Zahran Md Khudzari ,
Cheong Ping Pau, Mohd Faizal Ramli, Norfazlina Jaffar,
and Intan Fariza Gaaffar

Abstract Patient length of stay (LOS) in ICU and hospital's general care unit is one of the important indicators that is widely measured and benchmarked to improve the quality and efficiency of patient care. There are many studies both on statistical testing of the LOS outcome to determine factors associated with it and the outcome predictive modeling using machine learning algorithms. However, there are still fewer studies of the LOS outcome predictive modeling using local datasets. Therefore, an initial study of assessing supervised machine learning approaches for regression and classification tasks on predicting the LOS outcome has been conducted using an aggregated Multiparameter Intelligent Monitoring in Intensive Care-III (MIMIC-III) public dataset to obtain an overview of the prediction framework and to compare

F. Mahmud (✉)

Faculty of Electrical and Electronic Engineering, Microelectronics and Nanotechnology Shamsuddin Research Centre (MiNT-SRC), Institute of Integrated Engineering (I2E), Universiti Tun Hussein Onn Malaysia (UTHM), 86400 Parit Raja, Batu Pahat, Johor, Malaysia
e-mail: farhanah@uthm.edu.my

A. Z. M. Khudzari

School of Biomedical Engineering and Health Sciences, Faculty of Engineering, IJN-UTM Cardiovascular Engineering Centre (IJN-UTM CEC), Institute of Human Centered Engineering (iHUMEN), Universiti Teknologi Malaysia (UTM), 81310, Johor Bahru, Johor, Malaysia
e-mail: zahran.kl@utm.my

C. P. Pau

Cardiothoracic Surgery Department, Institut Jantung Negara (IJN), 145, Jalan Tun Razak, 50400 Kuala Lumpur, Wilayah Persekutuan Kuala Lumpur, Malaysia
e-mail: dr.paucp@ijn.com.my

M. F. Ramli · N. Jaffar · I. F. Gaaffar

Clinical Research Department (CRD), Institut Jantung Negara (IJN), 145, Jalan Tun Razak, 50400 Kuala Lumpur, Wilayah Persekutuan Kuala Lumpur, Malaysia
e-mail: faizal@ijn.com.my

N. Jaffar

e-mail: norfazlina@ijn.com.my

I. F. Gaaffar

e-mail: intanfariza_gaaffar@ijn.com.my

the predictive performance of the machine learning models. This is as a preparation for developing an outcome calculator of the LOS outcome and other operative outcomes in patients based on local data. The LOS was predicted using 10 classification and 15 regression models assessed using accuracy, precision, recall, and F1 scores for the classification task and root mean squared error (RMSE) for the regression task. The results showed that the Extreme Gradient Boosting classifier and Extreme Gradient Boosting regressor presented the highest validation and good testing performances for the LOS prediction. Although model overfitting trend was detected in both models' learning curves, in this respect, the study may serve as a useful starting point for an extended work associating with predictive modeling on the LOS and other operational outcomes.

32.1 Introduction

Patient length of stay (LOS) is one of the important indicators used in medical services to assess the efficiency of hospital management and patient quality of care. While, shorter LOS could reduce the burden of medical fees and increase the bed turnover rate, decreased LOS has been linked with improvements in treatment outcome and lower mortality rates, and with decreased infectiousness risk and medication side effects. Furthermore, the LOS in intensive care units (ICUs) is also a helpful parameter summarizing critical bed occupancy to enable optimization of resource allocation planning in ICUs and improve the performance of ICUs as the care provided by them is demanding, complex, and associated with high costs [1–4].

Models predicting the LOS outcome could play an important role in examining the efficiency of care in the hospitals, generally and in the ICUs, specifically. Many studies have been conducted on developing machine learning models for patient LOS prediction. According to a paper in [5], Tsai et al. had developed Artificial Neural Network (ANN) and linear regression models to predict LOS for cardiology patients at preadmission and predischarge stages and they found that both models could predict correctly for coronary atherosclerosis patients with high accuracy. Research conducted by Negassa et al. [6] has shown that a tree-structured prognostic classification model can be easily applied in predicting LOS among patients with elective percutaneous coronary intervention (PCI) to aid practitioners early on in their decision process. Moreover, a study had been conducted to predict ICU LOS from unplanned ICU admissions using eight regression models and it was found that it was difficult to predict the LOS of unplanned ICU admissions using patient characteristics at admission time only [7]. A study to predict the patient length of stay in ICUs after cardiac surgery using ANN and Adaptive Neuro-fuzzy System [8] shown result that the adaptive neuro-fuzzy algorithm could produce a more accurate model as it applies a hybrid algorithm from the neural networks and experts' knowledge, and it can identify nonlinear components. Meanwhile, the study using Multiparameter Intelligent Monitoring in Intensive Care-III (MIMIC-III) public dataset had been conducted by [9] to classify between short and long LOS for patients using neural networks which

had given a prediction accuracy of approximately 80% and outperformed a linear fit model which yielded an accuracy of 57%.

Moreover, discovering and validating a model using recent local data is recommended as healthcare and hospital policies differ between countries and over time [10]. However, there are relatively few studies of the operative outcome prediction models had been conducted locally in Malaysia. Thus, in this paper, we presented an initial study to predict the LOS operative outcome using several commonly used regression and classification types of supervised machine learning models with an aggregated MIMIC-III public dataset [11, 12] and to identify the best performance model for the prediction. The objective of this work is also to get an overview of the machine learning model construction framework to identify techniques in helping to achieve better performance of the LOS prediction. We aim to extend the pilot study using local data to develop an outcome calculator of the LOS outcome and other operative outcomes in patients.

32.2 Methodology

The pilot study has been conducted to assess supervised machine learning approaches for patient LOS prediction. It was done according to the flow of data science lifecycle [13] and Fig. 32.1 shows the study design flow with several processes that could be considered in each step. The study began with data pre-processing which contains sequential processes of obtaining data, data exploration, data cleaning and feature engineering. Only then, predictive modeling for the implementation of machine learning methods or artificial intelligence was done. The study was conducted through Python programming on Jupyter Notebook from the Anaconda data science platform.

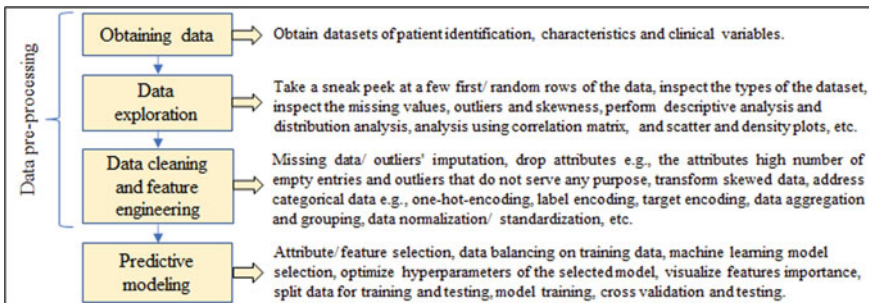


Fig. 32.1 Study design flow

32.2.1 Data Pre-Processing

In this study, the patient LOS has been predicted through supervised machine learning techniques with the aggregated public dataset of MIMIC-III [11, 12]. MIMIC-III is a database comprising de-identified health-related data associated with over forty thousand patients who stayed in critical care units of the Beth Israel Deaconess Medical Center between 2001 and 2012. The dataset has 58,976 rows with each row represents one admission and 28 attributes as summarized in Table 32.1.

Generally, before cleaning the data, data exploration needs to be considered to have a basic overall understanding of the data. This was done by taking a sneak peek at a few first or random rows of data and inspecting the types of the dataset, identifying missing values, outliers, and skewness, and conducting a descriptive analysis of the numerical data. Distribution analysis using histogram charts for numerical data and bar charts for analysis of categorical variables were also done in the data exploration.

In this study, 19 input attributes have been considered as the input variables which consist of the demographic data of gender and age, all the aggregated daily average numbers from several clinical data and the total number of daily interactions between the patient and the hospital that represents the summary of all the daily average numbers. Variables of text data with a lot of parameters, variables with missing data or that are irrelevant for the prediction have not been considered in this study. The conducted pre-processing techniques were one-hot-encoding and data normalization. The one-hot-encoding was conducted to the categorical-type of input attributes to convert them into numerical values which gave the final total number of 35 input attributes. Later, the data normalization was done to the input attributes to improve the performance of machine learning algorithms because some machine learning algorithms are sensitive to feature scaling.

Table 32.1 Attributes in the MIMIC-III dataset [11]

	Attributes	Descriptions
1	'hadm_id', 'gender', 'age', 'religion', 'marital_status', 'ethnicity'	Patient demographics
2	'insurance', 'admit_type', 'admit_location', 'AdmitDiagnosis', 'AdmitProcedure', 'NumCallouts', 'NumDiagnosis', 'NumProcs', 'NumCPTevents', 'NumInput', 'NumLabs', 'NumMicroLabs', 'NumNotes', 'NumOutput', 'NumRx', 'NumProcEvents', 'NumTransfers', 'NumChartEvents', 'ExpiredHospital', 'TotalNumInteract'	Admit information, daily average numbers of callouts for consultation, diagnoses, procedures, Current Procedural Terminology (CPT) events, input events, labs, microbiology labs, notes, output events, medical prescription, procedural events (insertion of an arterial line), transfers between care units, chart events, mortality, the summary of all the daily average numbers
3	'LOSdays', 'LOSgroupNum'	Patient length of stay in numerical and categorical data

32.2.2 Predictive Modeling

This study considered ten classification models known as Logistic regression, K-Nearest Neighbours, Gaussian Naive Bayes, Support Vector Machine, Decision Tree, Random Forest, Stochastic Gradient Descent, Gradient Boosting, Extreme Gradient Boosting and ANN classifiers and 15 regression models; Linear regression, K-Nearest Neighbours, Ridge, Least Absolute Shrinkage and Selection Operator (LASSO), Decision Tree, Random Forest, Extra Trees, Gradient Boosting, Extreme Gradient Boosting, Support Vector Regression (SVR), Linear SVR, ElasticNet, Bayesian Ridge, Kernel Ridge and ANN regressors [14, 15]. The predictive modeling was conducted for both classification and regression tasks to get an overview of the LOS prediction framework in both tasks using Python packages from Scikit-Learn, XGBoost and TensorFlow Keras [16–18].

For the ANN model, it was designed in a fully connected network structure with 35 input and four outputs for the classification task or one output for the regression task. One to six hidden layers of 71 hidden nodes were analyzed to obtain the optimum model design. The recommended number of hidden nodes is according to $n/2$, n , $2n$, and $2n + 1$ where n is the number of input nodes [19], and each input and output usually consist of one layer and hidden layer could consist of a minimum of one layer [20]. The ANN models were trained according to a gradient descent algorithm for 100 epochs. Rectified Linear Unit (ReLU) activation function was used in each hidden layer of the ANN classification and regression models, while softmax activation function was used in the output layer of the ANN classification model.

The prediction performance of the models in predicting the LOS outcome was according to RMSE for the regression task and accuracy, precision, recall, and F1 scores for the classification task. Here, the classification task was to classify four classes of the hospital LOS. 80% and 20% of the total dataset were used for the training and testing of the models, respectively. All the classification and regression models (except the ANN model) were trained with tenfold cross-validation using the training dataset and then, the model with the highest performance score in each task was further analyzed on its learning curves, and for its hyperparameter tuning and feature importance.

32.3 Results

Table 32.2 shows the summary of average validation performance scores of classification machine learning models with default parameter values based on tenfold cross-validation for the classification task in predicting the patient LOS, except for the ANN model which was based on the validation performance score after 100 epochs of training. While Table 32.3 presents the average RMSE score of regression machine learning models for the LOS prediction from the validation datasets of tenfold cross-validation, except for the ANN model which was based on the RMSE score after

Table 32.2 Summary of the validation performance scores of ten classification machine learning models for the LOS prediction

Classification model	Accuracy	Precision	Recall	F1-score
Logistic regression	0.737	0.703	0.689	0.685
K-Nearest Neighbours	0.659	0.630	0.621	0.623
Gaussian Naive Bayes	0.44	0.555	0.498	0.455
Support Vector Machine	0.727	0.692	0.668	0.657
Decision Tree	0.867	0.853	0.854	0.853
Random Forest	0.904	0.899	0.885	0.891
Stochastic Gradient Descent	0.658	0.572	0.612	0.571
Gradient Boosting	0.908	0.903	0.891	0.896
<i>Extreme Gradient Boosting</i>	<i>0.936</i>	<i>0.930</i>	<i>0.924</i>	<i>0.927</i>
ANN	0.831	0.827	0.799	0.807

Table 32.3 Summary of the validation performance scores of 15 regression machine learning models for the LOS prediction

Regression model	RMSE
Linear regression	11.74
K-Nearest Neighbours	9.44
Ridge	11.74
LASSO	12.15
Decision Tree	5.37
Random Forest	3.61
Extra Trees	3.59
Gradient Boosting	4.08
<i>Extreme Gradient Boosting</i>	<i>3.47</i>
SVR	10.88
Linear SVR	12.86
ElasticNet	12.06
Bayesian Ridge	11.74
Kernal Ridge	15.53
ANN	6.44

100 epochs of training. Based on the performance scores as shown in Table 32.2 and Table 32.3, the Extreme Gradient Boosting classifier and Extreme Gradient Boosting regressor (as marked in black) have given the highest performance. It is worth noting that ensembles of decision tree which are the Extreme Gradient Boosting, Random Forest, Gradient Boosting and Extra Trees have outperformed other types of machine learning models including the ANN in predicting the LOS outcome. While, between one to six hidden layers of the ANN model designs, the best performance scores of

the ANN model shown in Tables 32.2 and 32.3 were from the model designs with two and five hidden layers, respectively.

Besides, hyperparameter tuning of the Extreme Gradient Boosting classifier and Extreme Gradient Boosting regressor were furthermore conducted in effort to improve the performance of the models. The hyperparameter tuning was done according to common parameters for tree base learners which are the maximum depth per tree, learning rate, the number of trees in the ensemble and the fraction of columns to be randomly sampled for each tree [16]. Figure 32.2a and b show the learning curves of the Extreme Gradient Boosting classifier with default parameter values and after the hyperparameter tuning, while Fig. 32.2c and d show the learning curves of the Extreme Gradient Boosting regressor with default parameter values and after the hyperparameter tuning. The learning curves were inspected to analyze the error produced in training and validation stages by increasing the number of the training dataset. According to Fig. 32.2a–d, it was found in both models that, through a constant increase of the training sets size, the training error was already small from the beginning and approaching zero, but the error was much higher on the validation set. When the models were trained on more data, the validation error only decreased slightly which resulted in a quite large gap between the two learning curves in both

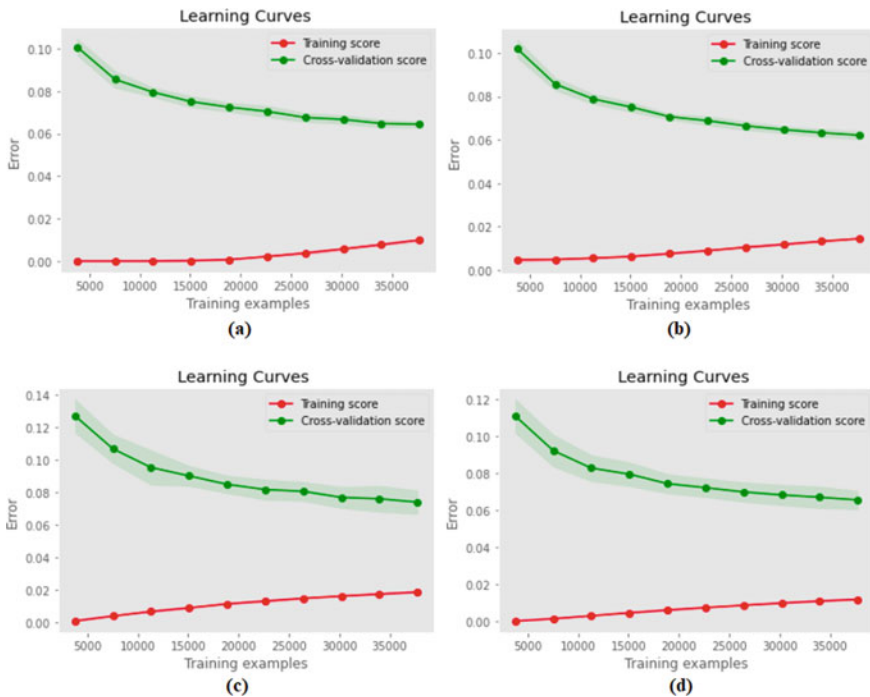


Fig. 32.2 The learning curves of the Extreme Gradient Boosting classifier with **a** default hyperparameter values and with **b** hyperparameter tuning. The learning curves of the Extreme Gradient Boosting regressor with **c** default hyperparameter values and with **d** hyperparameter tuning

Table 32.4 Summary of the testing performance scores of the Extreme Gradient Boosting classifier and the Extreme Gradient Boosting regressor for the LOS prediction. The numbers between parentheses are the scores from the models with default hyperparameter values

Machine learning model	Testing performance					
	Accuracy	Precision	Recall	F1-score	AUC ROC	RMSE
Extreme Gradient Boosting Classifier	0.940 (0.939)	0.936 (0.936)	0.930 (0.928)	0.933 (0.932)	0.955 (0.954)	X
Extreme Gradient Boosting regressor	X					3.244 (3.494)

models. This indicates the trend of overfitting in both models before and after the hyperparameter tuning. Thus, the models would benefit from more samples, feature engineering, or regularization of the learning algorithm to overcome the overfitting effect.

Nevertheless, most of the testing performance scores of the Extreme Gradient Boosting classifier and Extreme Gradient Boosting regressor as shown in Table 32.4 were slightly better than the validation performance scores in Table 32.2 and Table 32.3, respectively. Moreover, according to Table 32.4, the Extreme Gradient Boosting classifier and Extreme Gradient Boosting regressor with the hyperparameter tuning have a slightly more improvement in their prediction performance compared to the models with default hyperparameter values which are shown between parentheses. The Area Under the Curve of Receiver Characteristic Operator (AUC ROC) score for the classification prediction is also shown in Table 32.4.

In addition, based on the inspection of the feature importance score from the Extreme Gradient Boosting classifier and the Extreme Gradient Boosting regressor as shown in Fig. 32.3a and b, respectively, it was found that the ‘NumTransfers’ which represents the daily average number of transfers between care units has the highest feature importance score for the prediction using both models. The feature importance score is based on a mean decrease in node impurity which shows how useful it is at predicting the LOS outcome. Here, the mean decrease impurity is defined as the probability of reaching that node averaged over all trees of the ensemble [16].

32.4 Conclusion

Through this work, an initial study has been conducted on regression and classification tasks to predict the patient LOS using the aggregated MIMIC-III public dataset. As the result, the Extreme Gradient Boosting outperformed other machine learning

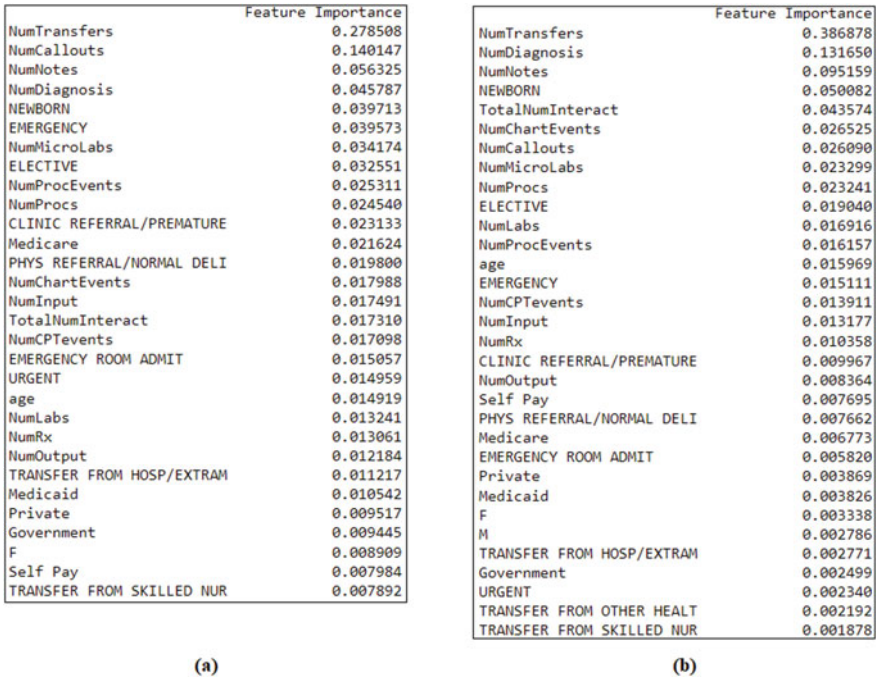


Fig. 32.3 The feature importance scores from **a** the Extreme Gradient Boosting classifier and **b** the Extreme Gradient Boosting regressor. The values shown are the mean decrease impurity more than 0.001

models in predicting the LOS operative outcome for both regression and classification tasks. Moreover, the k-fold validation training, hyperparameter tuning and learning curves plotting techniques applied during the machine learning modeling was helpful to achieve better model performance. However, the model overfitting trend interpreted from the validation and training learning curves of the models should be addressed carefully by understanding the data. This is to enable to apply better feature engineering apart from having a sufficient and good quality of data and regularizing the prediction model. Nonetheless, a comprehensive overview of the general framework for the LOS outcome prediction has been favorably obtained and several useful techniques involve in the prediction modeling have been identified. Therefore, we believe that this work will serve as a starting point to develop an outcome calculator of the LOS outcome and other operative outcomes in patients using a local dataset. This in turn will be able to be one of the applied intelligence solutions in providing an equitable and high-quality healthcare system in Malaysia.

Acknowledgements This work was a collaborative effort between UTHM, UTM and IJN. An acknowledgment is dedicated to the Faculty of Electrical and Electronic Engineering and MiNT-SRC, UTHM, IJN-UTM CEC, UTM, Cardiothoracic Surgery Department, IJN and CRD, IJN for a huge range of support for the research.









References

1. H. Baek, M. Cho, S. Kim, H. Hwang, M. Song, S. Yoo, Analysis of length of hospital stay using electronic health records: A statistical and data mining approach. *PloS one* **13**(4), p.e10195901 (2018)
2. H. Bueno, J.S. Ross, Y. Wang, J. Chen, M.T. Vidán, S.L.T. Normand, J.P. Curtis, E.E. Drye, J.H. Lichtman, P.S. Keenan, M. Kosiborod, Trends in length of stay and short-term outcomes among Medicare patients hospitalized for heart failure, 1993–2006. *JAMA* **303**(21), 2141–2147 (2010)
3. T. Rotter, L. Kinsman, E.L. James, A. Machotta, H. Gothe, J. Willis, P. Snow, J. Kugler, Clinical pathways: effects on professional practice, patient outcomes, length of stay and hospital costs, in *Cochrane database of systematic reviews* (3) (2010)
4. N. Lapidus, X. Zhou, F. Carrat, B. Riou, Y. Zhao, G. Hejblum, Biased and unbiased estimation of the average length of stay in intensive care units in the Covid-19 pandemic. *Ann. Intensive Care* **10**(1), 1–9 (2020)
5. P.F.J. Tsai, P.C. Chen, Y.Y. Chen, H.Y. Song, H.M. Lin, F.M. Lin, Q.P. Huang, Length of hospital stay prediction at the admission stage for cardiology patients using artificial neural network. *J. Healthc. Eng.* (2016)
6. A. Negassa, E.S. Monrad, Prediction of length of stay following elective percutaneous coronary intervention. *Int. Sch. Res. Not.* (2011)
7. I.W. Verburg, N.F. de Keizer, E. de Jonge, N. Peek, Comparison of regression methods for modeling intensive care length of stay. *PloS one* **9**(10), p.e109684 (2014)
8. H. Maharlou, S.R.N. Kalhori, S. Shahbazi, R. Ravangard, Predicting length of stay in intensive care units after cardiac surgery: comparison of artificial neural networks and adaptive neuro-fuzzy system. *Healthc. Inform. Res.* **24**(2), 109–117 (2018)
9. T. Gentimis, A.J. Alnaser, A. Durante, K. Cook, R. Steele, Predicting hospital length of stay using neural networks. *Int. J. Big Data Intell.* **6**(3–4), 297–306 (2019)
10. I.W.M. Verburg, A. Atashi, S. Eslami, R. Holman, A. Abu-Hanna, E. de Jonge, N. Peek, N.F. de Keizer, Which models can I use to predict adult ICU length of stay? A systematic review. *Crit. Care Med.* **45**(2), e222–e231 (2017)
11. M.D. Alexander Scarlat, <https://www.kaggle.com/drscarlat/mimic3d>. Last accessed 15 Feb 2021
12. A.E.W. Johnson, T.J. Pollard, L. Shen, L.H. Lehman, M. Feng, M. Ghassemi, B. Moody, P. Szolovits, L.A. Celi, R.G. Mark, MIMIC-III, a freely accessible critical care database. *Sci. Data* **3**(1), 1–9 (2016)
13. C.H. Lau, 5 Steps of a Data Science Project Lifecycle, 2019 Jan 3. <https://towardsdatascience.com/5-steps-of-a-data-science-project-lifecycle-26c50372b492>. Last accessed 20 May 2021
14. A.C. Müller, S. Guido, Introduction to machine learning with Python: a guide for data scientists, 1st edn. O'Reilly Media, Inc., 1005 Gravenstein Highway North, Sebastopol, CA 95472 (2016)
15. A. Singh, N. Thakur, A. Sharma, A review of supervised machine learning algorithms, in 2016 3rd International conference on computing for sustainable global development (INDIACom) (IEEE, 2016), pp 1310–1315
16. Pedregosa et al., Scikit-learn: machine learning in python. *JMLR* **12**, 2825–2830 (2011)
17. XGBoost Homepage, <https://xgboost.readthedocs.io/en/latest/python/index.html>. Last accessed 8 Oct 2021
18. Keras 2.0.0 Documentation Homepage, <https://faroit.com/keras-docs/2.0.0/>. Last accessed 8 Oct 2021
19. G. Zhang, B.E. Patuwo, M.Y. Hu, Forecasting with artificial neural networks: the state of the art. *Int. J. Forecast.* **14**(1), 35–62 (1998)
20. M.H.A. Yazid, M.H. Satria, S. Talib, N. Azman, Artificial neural network parameter tuning framework for heart disease classification, in 2018 5th international conference on electrical engineering, computer science and informatics (EECSI) (IEEE, 2018), pp. 674–679

Chapter 33

Prediction Model of Mass Rapid Transit Noise Level Using the Stepwise Regression Analysis



Nor Baizura Hamid , Mohd Erwan Sanik , Hafsa Mohammad Noor ,
Joewono Prasetyo , Mardiha Mokhtar ,
Mohamad Azim Mohammad Azmi , Mohamad Irwan Yahaya ,
and Mohd Zakwan Ramli 

Abstract Nowadays, rail transport is one of the most important transport modes chosen by Malaysians. However, the noise pollution caused by the railway causes complaints from residents living near this track. Therefore, the operator needs to order their workers to conduct monthly observations and measurements of their train noise level in the selected area. The conventional method requires more time and energy as the number of areas to monitor is various and the sound level measurement tools used are also expensive. Thus, a study was carried out to determine the current noise level

N. B. Hamid (✉) · M. E. Sanik · H. M. Noor · M. Mokhtar · M. A. M. Azmi
Spatial Technology for Civil Engineering (STforCE), Centre for Diploma Studies (CeDS),
Universiti Tun Hussein Onn Malaysia, Pagoh Education Hub, Pagoh, Johor, Malaysia
e-mail: norbaizura@uthm.edu.my

M. E. Sanik
e-mail: erwans@uthm.edu.my

H. M. Noor
e-mail: hafsa@uthm.edu.my

M. Mokhtar
e-mail: mardiha@uthm.edu.my

M. A. M. Azmi
e-mail: azim@uthm.edu.my

J. Prasetyo
Industry Centre of Excellence Railway (ICOE REL), University Tun Hussein Onn Malaysia,
Pagoh Education Hub, Pagoh, Johor, Malaysia
e-mail: joewono@uthm.edu.my

M. I. Yahaya
Faculty of Mechanical Engineering, Penang Branch, Universiti Teknologi Mara, Permatang Pauh,
Pulau Pinang, Malaysia
e-mail: irwan352@uitm.edu.my

M. Z. Ramli
College of Engineering, Universiti Tenaga Nasional, Kajang, Selangor, Malaysia
e-mail: zakwan@uniten.edu.my

produced by MRT train in the residential areas near to the Pusat Bandar Damansara station. The noise level measurement was conducted at Lorong Kasah Tepi and Medan Damansara Carpark, which are located nearby the Sungai Buloh—Kajang MRT Line. The noise level was measured at each location with three different slope distances using a sound level meter during operating and non-operating hours. Other than that, MRT speed and wind speed were measured as predictors to develop the Mass Rapid Transit Noise prediction model using the stepwise regression analysis. From the analysis, 88.37% of variation in Mass Rapid Transit Noise can be explained by the regression model.

33.1 Introduction

Noise is one of the causes of the deterioration of human health. Among the most notable effects of noise are disturbances that cause sleep deprivation [1]. In addition to sleep disturbances, there are previous studies showing that prolonged exposure to train noise also activates muscles during disturbed sleep as well as long-term health deterioration [2–4]. The sound intensity level on humans has been causing hearing loss particularly the people around train station [5]. Gradual hearing loss occurs due to continuous noise exposure. Sources of noise at train stations and nearby areas may be due to train engines, tire friction with rails, loudspeaker noise, and train horns [5].

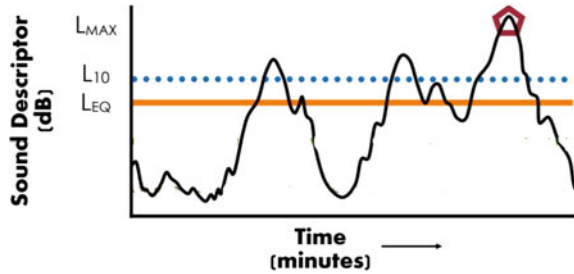
Noise prediction models, especially those related to land traffic have been the focus of research in recent decades. It is said to be a useful aid for the purpose of planning and upgrading systems and infrastructure such as railways. Therefore, regression analysis models are often used to leverage statistics to predict outcome values based on relevant variables. In addition, predictive models are also referred in the evaluation of existing or expected changes in noise conditions [6].

In this paper, the related works of developing prediction model is discussed in the next section. The objective of this study was to develop a prediction model of the Mass Rapid Transit (MRT) noise level by using several predictors that will be mentioned in Methodology. Then, the outcome of this study is presented and discussed in the Result and Discussion section. Lastly, this study will be concluded in the last section before acknowledgement and references list.

33.2 Related Works

Generally, noise prediction models are usually predicted using noise descriptors such as L_{EQ} , L_{10} , and L_{MAX} . The L_{EQ} also known as the Time-Equivalent Sound Level takes into account average sound fluctuations during louder and quieter moments, with more weightage to the louder, representing the equivalent continuous sound pressure level over a given period of time [7]. The L_{10} is a statistical descriptor of the sound level exceeded for 10% of the measurement period time, usually about

Fig. 33.1 Visual comparison of sound level represented by L_{EQ} , L_{10} and L_{MAX} (FHWA, 2017)



3 dB(A) greater than the L_{EQ} . The L_{MAX} also known as the Maximum Sound Level is the most peak sound level measured during a single noise event such as a train pass by, in which the sound level changes value along the time. The L_{MAX} is imperative in justifying the interference caused by a noise event with ambient sound [7]. Figure 33.1 shows the visual comparison of sound level represented by L_{EQ} , L_{10} and L_{MAX} . Sometimes, there is an 'A' stated in the descriptors such as $L_{A,EQ}$ that are meant to be referred as sound level for A-weighted measurement. A-weighted refers to a frequency dependent curve which is applied to sound pressure microphone measurements to impersonate the effects of human hearing [8].

In general, noise prediction models are used to assist noise measurement activities. It is more suitable for use when field measurements are difficult to carry out. Although field measurements can be made, it is possible that a certain maximum noise level in question can rarely be recorded. In addition to the difficulty of access to the receptor location, ambient noise may interfere with the accuracy of measurements in the field. Therefore, the noise prediction model becomes an option for the mentioned conditions [9].

Rail noise prediction models allow for the optimal management of noise prevention and mitigation to be considered by different scenarios [8]. The researcher introduced an equation that enables computation of $L_{A,MAX}$ for the Tehran-Karaj commuter train. The algorithm for predicting $L_{A,MAX}$, Y for this study was based on 50 measurements at various distances from 25 to 65 m from the middle of the track and at the height of 1.5 m. A linear regression equation was developed, as shown in Eq. (33.1).

$$Y = 0.999X + 0.080 \quad (33.1)$$

For Eq. (33.1), there was 89.8% located points on the regression line and indicates a reasonable fitness of the developed model for the field measurements [10].

In another study, there were two models developed namely the Railway Rolling Noise Prediction Software and the Track–Wheel Interaction Noise Software for predicting the sound pressure levels during a train passage due to several variables such as wheel/rail roughness, vibration dynamics, contact mechanics and sound radiation modules [11]. The field validation had been done under the European and Japanese conditions. However, this study focuses on the best method of reducing

noise by controlling it at its source. The information acquired from this study was an increase in railway rolling noise levels as train speed increases. Once the normal load is increased, the overall effect was a small reduction in the noise level. The effect of varying the stiffness of the vertical ballast was negligible. However, adjusting the vertical rail pad stiffness affects the noise radiation significantly. The noise produced by the rail decreases with an increase in stiffness, as the noise from the sleepers increases. Besides, increasing the factor of rail cross absorption will increase the noise produced by the train, thereby increasing the overall noise levels [11].

Another study was conducted in Hamadan city with the ultimate goal of developing a traffic noise model based on traffic conditions in Iranian cities [12]. A total of 282 samples were collected to develop a statistical regression model based on an A-weighted equivalent noise level for Iranian road conditions. The results revealed that the average $L_{A,EQ}$ in all stations was 69.04 ± 4.25 dB(A), the average speed of vehicles was 44.57 ± 11.46 km/h, and the average traffic load was 1231.9 ± 910.2 V/h with a high regression coefficient or R-square of 0.901 [12].

A study was conducted in developing a road traffic noise prediction model for the roads in Sri Lanka [13]. Researchers used the data of vehicle noise, vehicle class, vehicle speed, and the distance from the traffic flow line, which was obtained from several Western Province locations. Microsoft.Net[®] platform was used in the simulator and GUI development. The model's predictions were found to be within ± 11 dB(A) precision regarding the actual experimental observations [13].

In summary, the development of the noise prediction model for railway or road traffic is essential for the assessment of the environmental impact for future development. Therefore, noise can be minimized if the infrastructure is to be constructed. Also, the constant value of noise can be identified using the model when all predictors are omitted. This value can be compared with the permissible sound level recommended by the Department of Environment Malaysia, as shown in Table 33.1 [14]. In this study, based on the location of the data collection, the maximum permissible sound level is 60 dB(A) during day time and 50 dB(A) during night time.

33.3 Methodology

Two locations that are near the Pusat Bandar Damansara Station were selected in this study. These two locations labelled as A and B located at residential area and nearby the shophouses, respectively. Both locations from satellite images are as shown in Fig. 33.2.

The data was collected during normal working days. The total data point of noise $L_{A,MAX}$ for data collection was 120. Of these, ten measurements of MRT noise for every slope distance of 35, 40, and 50 m from the edge of the track were recorded. Different slope distances were chosen to show the pattern of data collected. The arrangement of the equipment for the measurement of train noise at different slope distance is illustrated in Fig. 33.3.

Table 33.1 Maximum permissible sound level (DOE, 2007)

Receiving land use category	Day time 7.00 am–10.00 pm [dB(A)]	Night time 10.00 pm–7.00 am [dB(A)]
Noise sensitive areas, low-density residential, institutional (school, hospital), worship areas	50	40
Suburban residential (medium density) areas, public spaces, parks, recreational areas	55	45
Urban residential (high density) areas, designated mixed development areas (residential-commercial)	60	50
Commercial business zones	65	55
Designated industrial	70	60

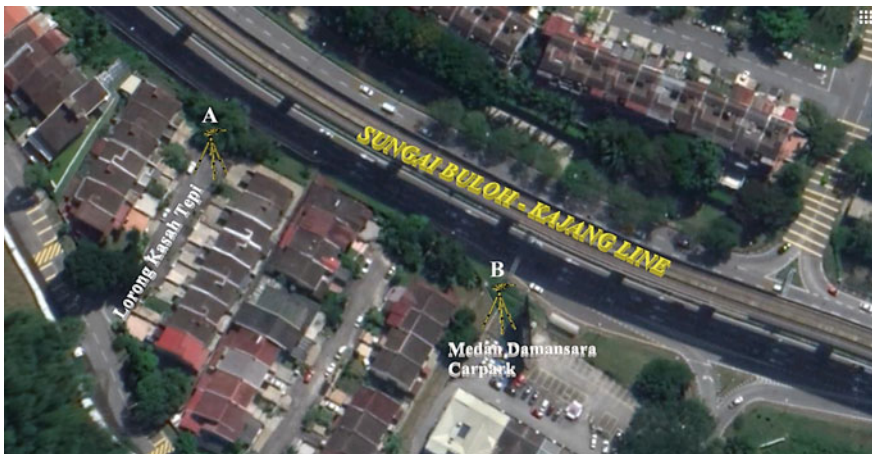


Fig. 33.2 Location A (Lorong Kasah Tepi) and Location B (Medan Damansara Carpark)

The MRT noise values, $L_{A, MAX}$ were measured and used as a dependent variable or response in the prediction model using stepwise regression analysis. Other data that were collected as independent variables or predictors are the speed of the MRT train, wind speed, and slope distance during the off-peak and peak hours. The MRT speed is estimated using Eq. (33.2).

$$Speed = \frac{Length\ of\ MRT\ train}{Time\ taken\ to\ pass\ by\ from\ head\ to\ tail\ at\ a\ point} \quad (33.2)$$

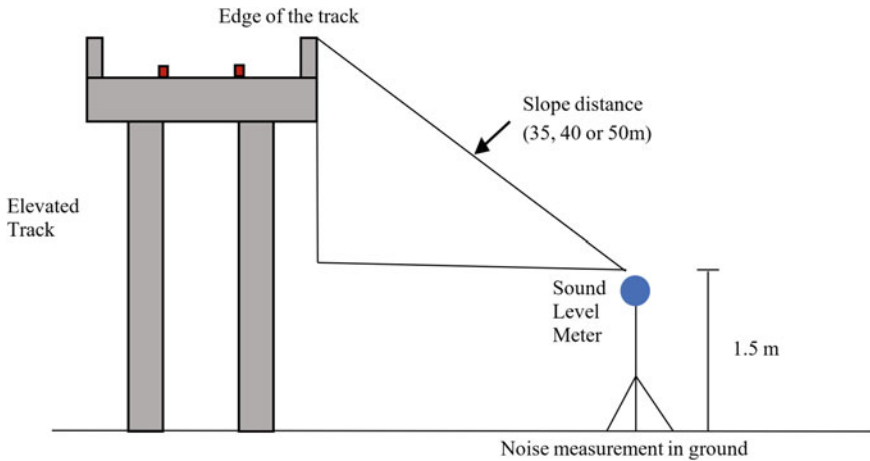


Fig. 33.3 The schematic diagram for field measurement set-up

A stopwatch is needed to record the time taken from the head of a train to pass by a point to the end of the train (tail). The total length of the MRT train is 89.56 m. Wind speed was measured using the Digital Anemometer MS6252A. The train operates from 7 to 9 am and 5 pm to 7 pm, considered as operating hours, while apart from both periods is named non-operating hours.

Few precautions were taken in data collection such as the windscreen should always be used during the measurement. Ambient noise must be low, and the sound level meter should be held at arm's length, usually mounted on a tripod, to minimize disturbance of the sound field due to the operator's body. The sound level meter should be at a height between 1.2 and 1.5 m from the ground and should have a distance of at least 0.5 m away from the body of the observer to avoid reflections, which can cause an error up to 6 dB(A) [15].

Prediction model was obtained using the stepwise fit regression model in the MINITAB 19 as the tool. The response or the dependent variable, Y is MRT noise level ($L_{A, MAX}$) and predictors or independent variables consist of MRT speed, wind speed, slope distance, and one categorical variable which is type of hours either non-operating (NOH) or operating hours (OH). This model fits two-way interactions and quadratic terms. If the model fits the data well, this developed equation can be used to predict MRT Noise for specific values of the X variables, or find the settings for the X variables that correspond to a desired value or range of values for MRT Noise.

33.4 Result and Discussion

Stepwise Regression analysis was carried out to develop the model. The following terms or predictors are in the fitted equation that models the relationship between Y and X variables:

- X₁: MRT Speed.
- X₂: Wind Speed.
- X₃: Slope Distance.
- X₄: NOH (0)/OH (1) (Categorical data).
- X₂².
- X₁*X₃.

Correlation analysis was carried out before the stepwise regression analysis to check correlation between Y and X₁, X₂, X₃ and X₄ variables. Table 33.2 shows the Pearson correlation values obtained in the analysis. The description of correlation was based on the study done by Miyamoto et al. [16].

Referring to Table 33.4, all independent variables were correlated to the dependent variable. Therefore, these variables were considered in the stepwise regression analysis. Categorical variable X₄ was not included in the model equation, however, the model was split into two separate equations representing the non-operating (NOH) and operating hours (OH) categories, respectively. While variables (e) and (f) were created due to the two-way interactions and quadratic terms fitted in the model. Figure 33.4 shows the model building using the stepwise regression analysis.

Referring to Fig. 33.4, the variations that were accordingly done to the model in order to predict Y were shown. Step P is referring to *p*-value obtained along the steps in the fitting of the terms in the model. While Final P is referring to the final *p*-value after all steps being completed. The stepwise regression procedure consists of two steps. Firstly, the variable with highest association to the dependent variable Y was picked. Next, any variables that can be added and removed in the current model were beheld to complete the procedure. Table 33.3 shows the details of regression analysis.

Referring to Table 33.3, in the stepwise process, regression model must adhere to the hierarchy rules which is if one of the higher-order variables is statistically significant, all the lower-order variables that create the higher-order variable shall be included although they are not statistically significant. In this case, since the interaction term of X₁*X₃ is statistically significant with *p*-value that is less than 0.05, both MRT speed, X₁ and slope distance, X₃ terms must be included in the

Table 33.2 Correlation Analysis

Variable	MRT Noise, Y	Description
MRT speed, X ₁	0.401	Correlated
Wind speed, X ₂	0.635	Correlated
Slope distance, X ₃	-0.615	Correlated
NOH/OH, X ₄	-0.673	Correlated

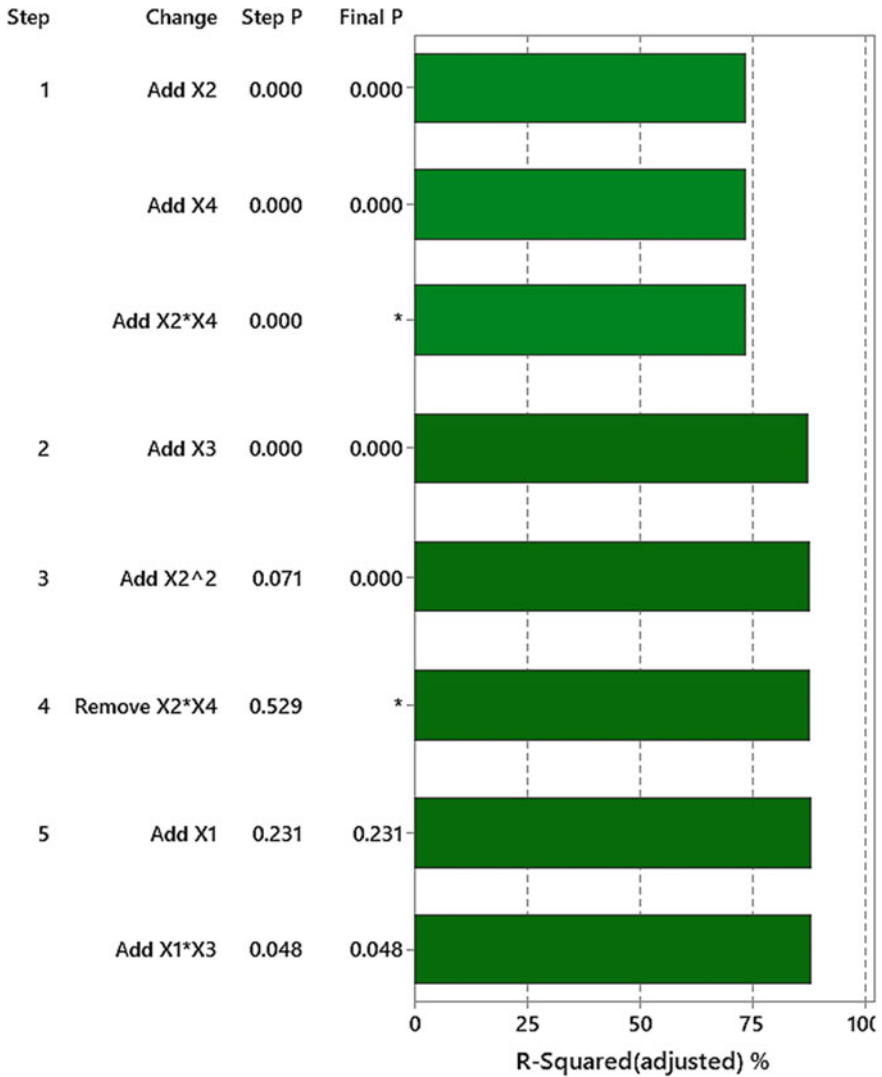


Fig. 33.4 Model building sequence

model. Therefore, the inclusion of both variables with p -values of greater than 0.05 were permitted in the model due to the rules of hierarchy. While, the wind speed term has p -value less than 0.05, thus, it is statistically significant at 95% significance. The consideration of quadratic terms resulted in the additional significant term of X_2^2 with p -value of less than 0.05. In the last line of Table 33.2, represented as '1' in term's column, the coefficient for operating hour (OH) estimates that the mean MRT speed decreases by 5.028 of constant value. While constant value of 51.3 is for the model during non-operating hours (NOH).

Table 33.3 Regression analysis result

Term	Coefficient	Standard error coefficient	T-Value	p-Value
Constant	51.30	7.65	6.71	0.000
MRT speed, X1	0.218	0.118	1.84	0.231
Wind speed, X2	24.68	4.09	6.03	0.000
Slope distance, X3	0.189	0.177	1.07	0.289
X2 ²	-11.27	1.78	-6.32	0.000
X1*X3	-0.00567	0.00284	-2.00	0.048
NOH/OH, X4				
1	-5.028	0.553	-9.10	0.000

Table 33.4 Model Summary

S	R-square	R-square (adjusted)	R-square (predicted)
0.78	88.37%	87.75%	86.98%

Table 33.4 shows the model summary in terms of standard error of regression (S), R-square, adjusted and predicted R-Squares. The model explains approximately 88% of the variation in the response. For these data, the R-square value indicates the model provides an adequate fit to the data.

The final model consists of two categories which are non-operating and operating hours as shown in Eqs. (33.3) and (33.4), respectively.

$$Y = 51.30 + 0.218X_1 + 24.68X_2 + 0.189X_3 - 11.27X_2^2 - 0.00567X_1X_3 \tag{33.3}$$

$$Y = 46.27 + 0.218X_1 + 24.68X_2 + 0.189X_3 - 11.27X_2^2 - 0.00567X_1X_3 \tag{33.4}$$

Referring to Table 33.4, the adjusted R-Square value for both equations is 0.8775, which shows that 87.75% or 88% of variation in Y can be explained by the regression model. While, approximately 95% of the data points lie between the regression line and ± 1.56 dBA of MRT noise level. Referring to Table 33.1 and the obtained models, the MRT noise is still below the permissible sound level if all predictors are omitted. The range of independent variables or predictors is as shown in Table 33.5. The final model can be accurately applied if the values of predictors are in the range.

Table 33.5 Range of values for independent variables or predictors in the developed model

Predictors	Range
Train speed, X ₁	53.7–64.5 km/h
Wind speed, X ₂	0.8–1.6 m/s
Slope distance, X ₃	35–50 m

33.5 Conclusion and Recommendation

In conclusion, the prediction model for the MRT noise was obtained using MRT speed, wind speed, and slope distance during non-operating and operating hours as independent variables. The stepwise regression analysis was used in the development of the model since it has been used in the previous studies and a good percentage of variation in dependent variables can be explained by the model. The model can be referred to Eqs. (33.3) and (33.4) in the previous section. The model would be useful for the MRT operators to monitor the noise level from time to time based on the predictors' values. Since 88% variation in MRT noise can be explained by the model, there is still another 12% that contributed by unknown variables that need to be studied in the future. Besides that, accuracy of data can be improved by using proper equipment such as speed laser gun and taking into consideration the presence of ambient noise in the measurement as well.

Acknowledgements The authors would like to thank the Research Management Centre, Universiti Tun Hussein Onn Malaysia for the fund given under the Tier 1 Research Grant Code H813.

References

1. R. Khaiwal, T. Singh, J.P. Tripathy, S. Mor, S. Munjal, B. Patro, N. Panda, Assessment of noise pollution in and around a sensitive zone in North India and its non-auditory impacts. *Sci. Total Environ.* **566–567**, 981–987 (2016)
2. S. Shahidan, N.M. Bunnori, N. Md Nor, S.R. Basri, Damage severity evaluation on reinforced concrete beam by means of acoustic emission signal and intensity analysis. *IEEE Sympos. Indus. Electron. Appl.* **2011**, 337–341 (2011)
3. G.J.L. Micheli, S. Farné, Urban railway traffic noise: looking for the minimum cost for the whole community. *Appl. Acoust.* **113**, 121–131 (2016)
4. N.I.R. Ramzi Hannan, S. Shahidan, M.Z. Maarof, N. Ali, Physical and chemical properties of coal bottom ash (CBA) from Tanjung Bin Power Plant. *IOP Conf. Ser. Mater. Sci. Eng.* **160**, 12056 (2016)
5. M.Y. Indrayani, A. Amiruddin, B. Arnita, Analysis of train noise level at Bandar Khalipah station, Deli Serdang using sound level meter 130 dB. *J. Phys: Conf. Ser.* **1811**, 012017 (2021). <https://doi.org/10.1088/1742-6596/1811/1/012017>
6. R. Golmohammadi, A compact model for predicting road traffic noise, Iran. *J. Environ. Health Sci. Eng.* **6**(3), 181–186 (2009)
7. Federal Highway Administration (FHWA), U.S. Department of Transportation. Highway traffic noise resources, sound level descriptors-FHWA-HEP-17-053, May 2017
8. A.M.O. Mohamed, E.K. Paleologos, F.M. Howari, Noise pollution and its impact on human health and the environment, Chapter 19, in *Pollution Assessment for Sustainable Practices in Applied Sciences and Engineering*, ed. by A.M.O. Mohamed, E.K. Paleologos, F.M. Howari (Butterworth-Heinemann, Oxford, 2021), pp. 975–1026
9. N. Hassan, A study of usability noise prediction model by using linear regression analysis for KTMB commuter between Shah Alam to Bukit Badak Route, Universiti of Tun Hussein Onn Malaysia (2016)
10. P. Nassiri, M. Abbaspour, M. Mahmoodi, S. Givargis, A rail noise prediction model for the Tehran-Karaj commuter train. *Appl. Acoust.* **68**(3), 326–333 (2007)

11. S. Jiang, P.A. Meehan, D.J. Thompson, C.J.C. Jones, Railway rolling noise prediction: field validation and sensitivity analysis. *Inter. J. Rail Transp.* **1**(1–2), 109–127 (2013). <https://doi.org/10.1080/23248378.2013.788359>
12. R. Golmohammadi, M. Abbaspour, P. Nassiri, H. Mahjub, A compact model for predicting road traffic noise, Iran. *J. Environ. Health. Sci. Eng.* **6**(3), 181–186 (2009)
13. R.T. Sooriyaarachchi, D.U.J. Sonnadara, Development of a road traffic noise prediction model. *Proceedings of the technical session*, vol. 22 (Institute of Physics, Sri Lanka, 2006) pp. 17–24
14. Department of Environment, *The planning guidelines for environmental noise limits and control, book 1 of 3* (2007)
15. N. Kumar, Study of vehicular traffic noise and its prediction. Master Thesis. Thapar University, Patiala-147004, 2009
16. A. Miyamoto, J. Puttonen, A. Yabe, Long term application of a vehicle-based health monitoring system to short and medium span bridges and damage detection sensitivity. *Engineering* **09**, 68–122 (2017)

Chapter 34

Engine Monitoring During Tuning Process for Diesel-Compressed Natural Gas (CNG) Dual Fuel Engine Using Statistical Approach



Abd Fathul Hakim Zulkiffi , Mas Fawzi Mohd Ali ,
Muammar Mukhsin Ismail, and Shahrul Azmir Osman 

Abstract Engine monitoring during tuning process is vital for diesel-Compressed Natural Gas (CNG) dual fuel (DDF) engine to ensure the dual fuel operation run without any abnormalities. Knock sensor which is vibro-acoustic sensor based is commonly used for engine monitoring. However, high variation of vibration signal due to different background noises can leads to faults detection. The objective of this paper is to introduce the statistical method to overcome this issue. A diesel engine converted into DDF system was investigated at different dual fuel ratio, which is based on CNG substitution rates (0–40%) and engine speed in a range of 1500–3500 rpm. The DDF engine was monitored using commercial knock sensor, which is mounted at the outer wall of the engine. The frequency distribution analysis (FDA) was applied to interpret the high variability of the vibro-acoustic signal. The results showed high CNG substitution rates can be achieved during intermediate to high engine speeds, while at low engine speed, maximum CNG substitutions rate is 10%. In addition, the FDA provides additional information regard to engine stability, which is graphically presented using the box-plot. Therefore, the functionality of the knock sensor can be enhanced despite it is originally used for knock monitoring.

A. F. H. Zulkiffi (✉)

Centre of Automotive and Powertrain Technology (CAPT), Faculty of Engineering Technology, Campus (Pagoh Branch), Pagoh Higher Educational Hub, KM 1, Jalan Panchor, 84600 Pagoh, Muar, Johor, Malaysia
e-mail: fathul@uthm.edu.my

M. F. M. Ali · M. M. Ismail · S. A. Osman

Automotive Research Research Group, Centre for Energy and Industrial Environment Studies (CEIES), Faculty of Mechanical and Manufacturing Engineering, Universiti Tun Hussein Onn Malaysia, Parit Raja, 86400 Batu Pahat, Johor, Malaysia

34.1 Introduction

Faulty diagnosis for engine monitoring during the diesel-CNG dual fuel (DDF) tuning process is crucial to avoid any abnormalities during dual fuels operation and to ensure their stability. A common method used during engine tuning is using aftermarket knock detectors such as Phormula and PLEX tuning, which are based on the vibro-acoustic knock detection method [1, 2]. Technically, the vibration generated by structural oscillations emitted by different sources within the engine, such as the combustion process (in-cylinder pressure oscillations) and the impact of mechanical components (piston slap, gear and impact of the intake/exhaust valve), as shown in Fig. 34.1. These vibrations produce noises that could be heard in the frequency area of human ears (20–20,000 Hz). It shows that engine monitoring faults detection is not easy due to different background noises from an engine.

Several established engine-monitoring methods have been found in the literature in the past two decades. For example, in order to improve the knock detection at high speed, the pattern classification based on cepstrum analysis from vibration input was proposed [3]. The decomposition of the signal was studied using independent component analysis (ICA) to break down the diesel engine vibration signal [4]. Klinchaeam presented a statistical analysis of the vibration signal based on an angular domain in order to determine the impact between valves and tappets and between the plug and the cylinder head [5]. Devasenapati also used statistical approach in time domain to detect misfire problems [6].

Delvecchio et al. studied the symmetrized dot pattern (SDP) for visual characterization in vibration analysis to diagnosis faulty during engine tested in motor condition [7, 8]. Other techniques based on time–frequency analysis, such as Short-Time

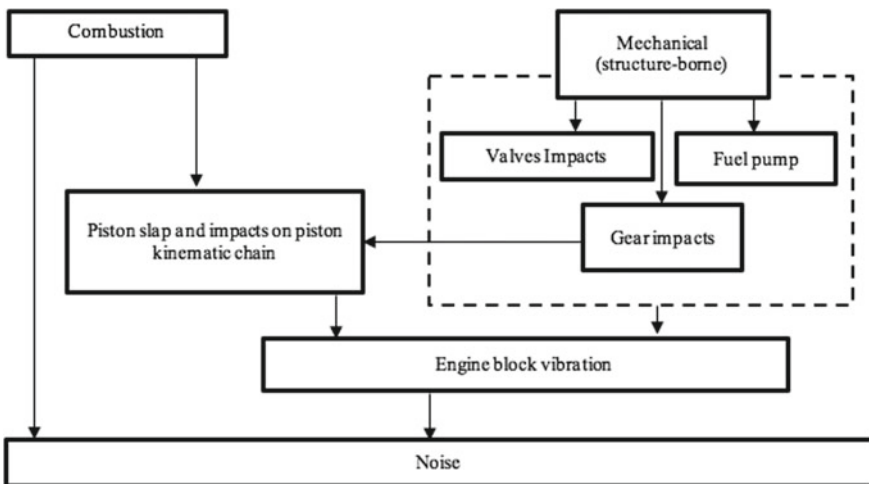


Fig. 34.1 Noise sources in internal combustion engine

Fourier Transform (STFT), Continuous Wavelet Transform (CWT) and Wigner-Ville Distribution (WVD), were also used in engine monitoring [9–11].

Antoni proposed cyclostationary approaches to various combustion defects caused by injection timing, misfire and knocking in vibration analysis [12, 13]. This technique was subsequently applied to engine cold test monitoring [14].

Thus, this research aims to introduce the statistical approach using frequency distribution analysis for engine monitoring purposes during tuning process. This method capable to overcome the unwanted different background noises that came from an engine. In addition, the engine stability can be established.

34.2 Research Methodology

34.2.1 Fuel Properties

For this research, the fuel resources are obtained from commercial refueling station. For Malaysia market, the diesel fuel is Euro 2 M standard which contained 7% of palm methyl ester (PME). This fuel is followed with the Malaysian standard MS 2008. Table 34.1 shows the fuel properties for diesel and CNG.

34.2.2 Experiment Setup

The layout for the experiment setup is shown in Fig. 34.2. Toyota Hilux 2.5L common rail diesel engine (2KD-FTV) was used as the test engine. The engine has been converted into diesel-CNG dual fuel systems using aftermarket dual fuel conversion kits. The conversion method was followed the Malaysian legislation as per discussed in [16]. The engine specification is shown in Table 34.2. The Dynapack 4WD chassis dynamometer (Model No: 4022) was used to conduct the steady-state dynamometer testing, which is followed the SAEJ1349 standard.

The engine was tested at constant speed for 1500–3500 rpm with 500 rpm interval. The dual fuel ratio was set based on CNG substitution rates, which are 0, 10, 20, 30 and 40%. The 0% substitution rate represent the diesel mode condition. The CNG flow rate was monitored using the ALICAT Scientific Mass Flow Meter (M-250SLPM), while the diesel flow rate was monitored using the ECU Diagnostics Bosch KTS 570. The fuel substitution rates were determined based on mass fuel ratio and followed as per presented in [17].

The commercial knock sensor was used to measure the vibro-acoustic signal from the test engine. The knock sensor was placed at the outer wall of the engine, which closed to the first cylinder where the pressure transducer is installed. Then, the vibro-acoustic signal is amplified using commercial audio amplifier before connected to data acquisition (cDAQ NI-9188). The DEWETRON encoder instrument (Model:

Table 34.1 Fuel properties of diesel and CNG

Properties	Diesel	CNG
Flash point (PM, °C)	76	–
Kinematic viscosity (40 °C, mm ² /s)	3.21	–
Sulfur (mg/kg)	7.5	–
Cetane index	52	–
Density (15 °C, kg/m ³)	831	–
Low heating value (MJ/kg)	43.15	–
Gross heating value (MJ/Sm ³)		39.20
Specific gravity (compare to air)		0.6042
Flammability limit (%)		5–15
Compressibility		0.9977
Methane (vol.%)		93.07
Ethane (vol.%)		3.70
Propane (vol.%)		0.90
i-Butane (vol.%)		0.29
n-Butane (vol.%)		0.13
i-Pentane (vol.%)		0.07
C6+ (vol.%)		0.07
Nitrogen (vol.%)		0.68
Carbon dioxide (vol.%)		1.10

Reproduced from [15]

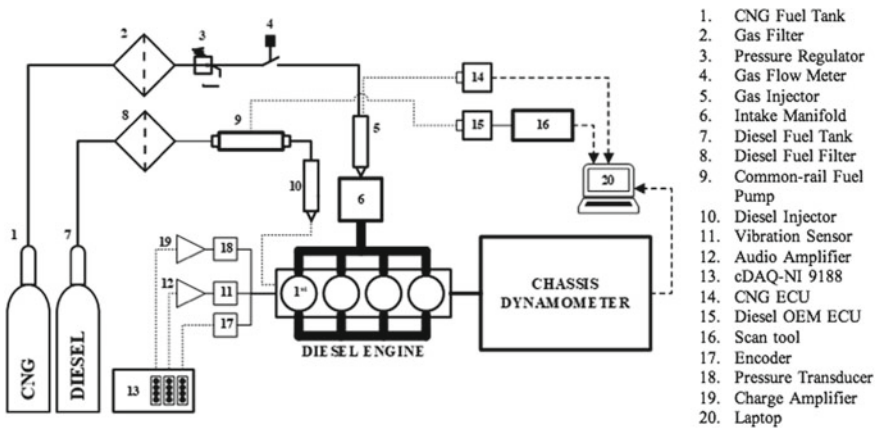


Fig. 34.2 Schematic diagram of the experiment setup

Table 34.2 Specification of test engine

Engine Specification	Descriptions
Engine code	2KD-FTV
Bore × stroke	92.0 × 93.8 mm
Engine displacement	2494 cc
Compression ratio	17.4:1
Fuel injection system	Common rail direct injection
Maximum power	80 kW @ 3600 rpm
Maximum torque	325 Nm @ 2000 rpm

CA-RIE 360) was used to generate the TDC and clock signals. The combustion pressure was measured using Kistler pressure transducer (type 6056a) and conditioned by Kistler charge amplifier (type 5018a).

34.2.3 Establishment of Acoustic Index

The vibro-acoustic signal was recorded in crank angle domain as shown in Fig. 34.3. Then, the vibro-acoustic signal from knock sensor was analyzed using BRIC signal processing method to establish the acoustic index. BRIC means band-pass, rectify, integrate and compare. It is a common method used for knock sensor application especially for vibro-acoustic signal analysis. Two calibration windows were defined, which are reference window and acoustic window. Each window was set at 30 CAD.

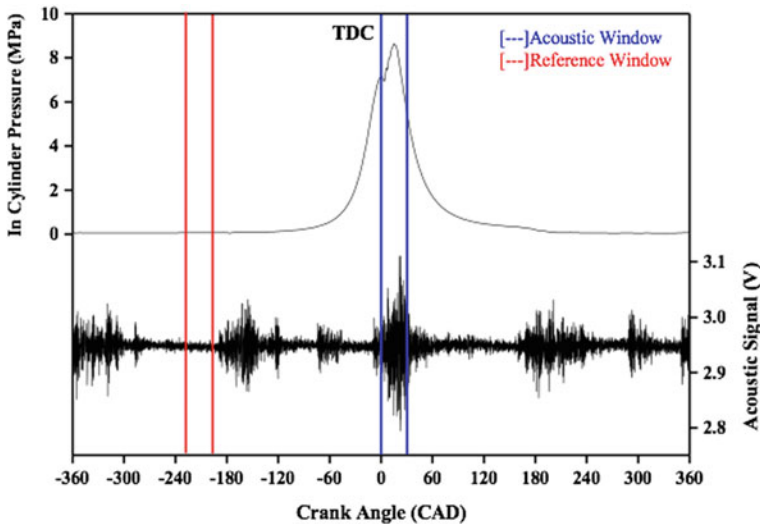


Fig. 34.3 The calibration window for BRIC signal processing method

The acoustic window is defined as the window of the signal before and after TDC, while the reference window is the window where there is a little noise. The draper acoustic pressure wave formula was used to determine the cut-off frequencies for band-pass filters as shown in Eq. (34.1) [18]. The in-cylinder temperature was calculated using ideal gas law equation as shown in Eq. (34.2) where P is the combustion pressure, V is the combustion chamber volume, m is the trapped mass for burned fuels (diesel and CNG) and air, and R is the gas constant. Therefore, the acoustic index before TDC is defined as AI BTDC and the acoustic index after TDC is determined as AI ATDC. For this research, both conditions before and after TDC were investigated to evaluate the vibration intensity before and after combustion process happen. This is required to understand the combustion stability of the diesel-CNG dual fuel system.

$$f_{u,s} = \frac{\rho_{u,s} \cdot a}{\pi \cdot B} = \frac{\rho_{u,s} \cdot \sqrt{kRT_{cyl}}}{\pi \cdot B} \quad (34.1)$$

where;

$f_{u,s}$	Oscillation frequency of in-cylinder gas
$\rho_{u,s}$	Vibration mode factor
u	The circumferential oscillation mode number
s	The radial oscillation mode number
a	Velocity of sound (m/s)
B	Cylinder Bore (mm)
k	Ratio of specific heat
R	The gas constant of working gas [J/(kg K)]
T_{cyl}	Absolute temperature of working gas [K] or the in-cylinder temperature

$$T_{cyl} = \frac{PV}{mR} \quad (34.2)$$

34.2.4 Uncertainty Analysis

The overall uncertainty (U_{overall}) for the experimental results is examined by combining systematic uncertainty (S_u) and random uncertainty (R_u) using the root-sum-square method (RSS) with a 95% confident level of true value [refer Eqs. (34.3)–(34.5)] [19–21].

Based on Eqs. (34.4) and (34.5), M is the physical parameters dependent on each variable, X_i . The S_i and R_i represent the uncertainty in M and the measuring range respectively. The overall uncertainty obtained for the experiment is $\pm 0.48\%$. The uncertainty for each measurement parameter is summarized in Table 34.3.

Table 34.3 Accuracy and uncertainty of measured parameters

Equipment/instrument	Measured parameter	Range	Accuracy	Uncertainty (%)
Chassis dynamometer	Engine torque		±1.84 Nm	±0.2
	Engine power		±0.48 kW	±0.05
Gas flow meter	CNG flow rate	0–250 splm	±0.2%	±0.08
Pressure transducer	In-cylinder pressure		±0.1 MPa	±0.05
Crank angle encoder	Crank angle degree	0–360°	±1°	±0.3
Vibration sensor	acoustic index		±0.64	±0.3

$$U_{overall} = \sqrt{S_u^2 + R_u^2} \quad (34.3)$$

$$\frac{S_u}{M} = \left[\sum_{i=1}^n \left(\frac{1}{M} \frac{\partial M}{\partial X_i} S_i \right)^2 \right]^{1/2} \quad (34.4)$$

$$\frac{R_u}{M} = \left[\sum_{i=1}^n \left(\frac{1}{M} \frac{\partial M}{\partial X_i} R_i \right)^2 \right]^{1/2} \quad (34.5)$$

34.3 Results and Discussions

The vibro-acoustic signal analysis was performed by offline mode where 200 consecutive cycles of acoustic index before Top Death Centre (AI BTDC), after Top Death Centre (AI ATDC), and peak combustion pressure (PP) was observed. The coefficient of variance (COV) of AI BTDC, AI ATDC and PP were calculated and presented in Fig. 34.4. The result showed high variability of acoustic index where 21.54% and 23.71% COV were recorded for AI BTDC and AI ATDC, respectively. It is contradicted to COV of PP where less than 1% was recorded. High COV means some of the acoustic index data are categorized as outliers. These outliers came from the different unwanted background noises, which are contributed to the error. This error will be more significant when the COV is increased.

The frequency distribution analysis was proposed to overcome this problem, where the box-plot is utilized to display the acoustic index data via its quartiles. Figure 34.5 shows the legend for box plot used in this research. The box indicated the degree of dispersion through interquartile range (IQR) for the data where the upper line is the first quartiles (Q1), and the lower line is third quartiles (Q3). The median line indicates the skewness of the data and the mean values shows an average data for IQR without the outliers.

Figure 34.6 shows the acoustic index range against CNG substitution rates for 1500 rpm. The vibration intensity at BTDC shows approximately stable condition

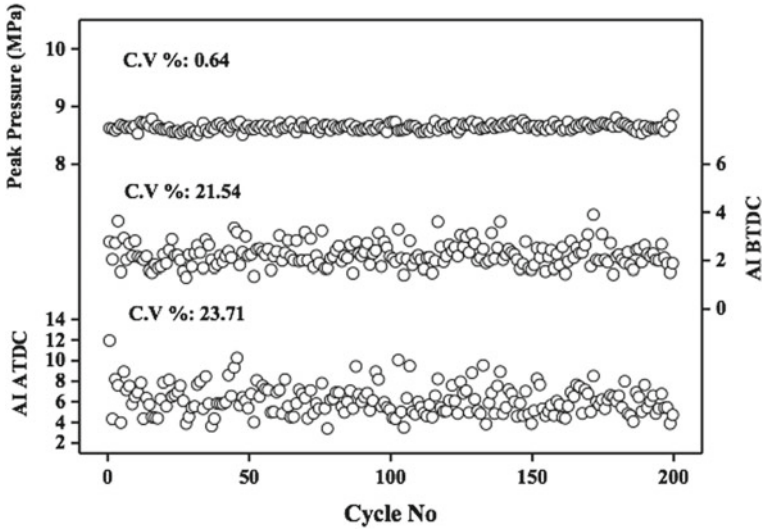
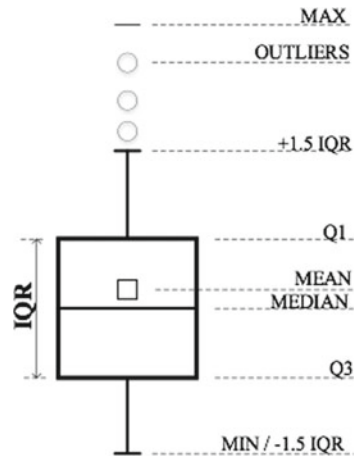


Fig. 34.4 Coefficient of variance (COV) for PP, AI BTDC and AI ATDC at 1500 rpm diesel operation

Fig. 34.5 Legend for the box-plot



even though running the engine with the dual fuel operation. High vibration intensity was observed at ATDC especially during diesel mode. The CNG addition was stabilized the engine where the vibration intensity decreases as CNG substitution rates increases. In addition, the box size for dual fuel mode were comparatively shorter compared to diesel mode, which indicates better engine stability due to lower variance of acoustic index.

Figure 34.7 shows the acoustic index range against CNG substitution rates at 2000 rpm. Both BTDC and ATDC conditions shows high CNG substitution increases the

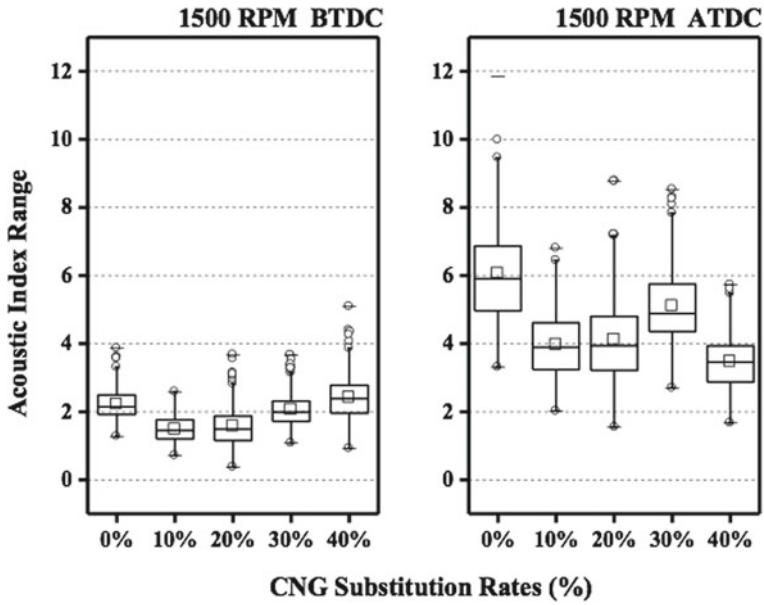


Fig. 34.6 Acoustic index versus CNG substitution rates for 1500 rpm

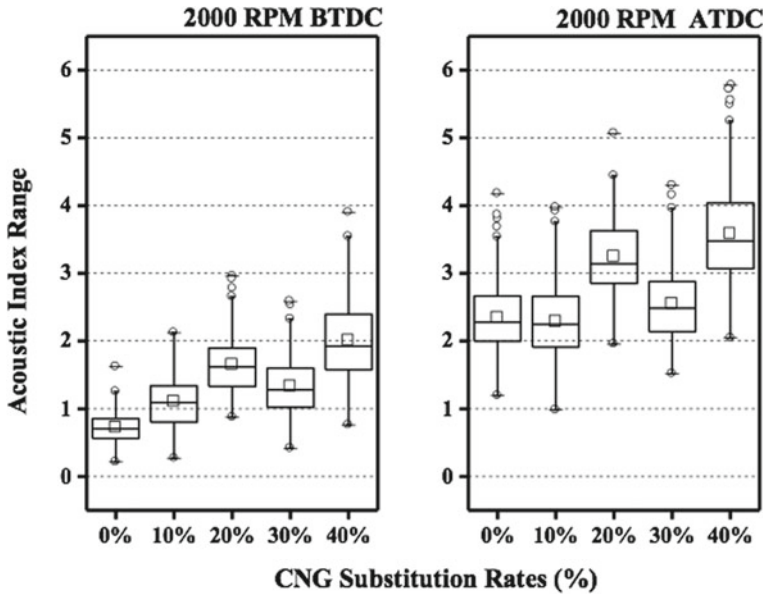


Fig. 34.7 Acoustic index versus CNG substitution rates for 2000 rpm

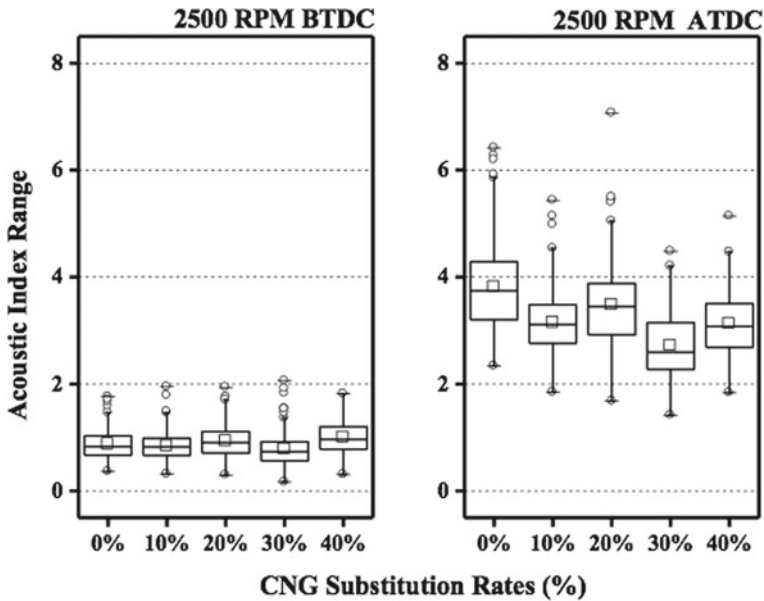


Fig. 34.8 Acoustic index versus CNG substitution rates for 2500 rpm

vibration intensity, except for 30% CNG substitution, where both conditions show considerably lower. However, for 2500 rpm, the engine shows stable condition even with high CNG substitutions rates (see Fig. 34.8). Similar findings were observed at 3000 rpm and 3500 rpm, as shown in Figs. 34.9 and 34.10, respectively. This supports by the box sizes where not much different between diesel and dual fuel mode. Its reveals that the test engine is capable to run dual fuel mode with high CNG substitution rates during intermediate to high engine speeds.

The frequency distribution analysis through box plot provides useful information where the vibration intensity from an engine could be further understand. The variance of acoustic index shows it is capable to measure the stability of the engine and delivers a good indicator during tuning process.

34.4 Conclusion

The statistical approach for engine monitoring during DDF tuning process was presented. Several conclusions can be drawn from the results obtained as follows:

1. The vibro-acoustic signal processing through the frequency distribution analysis was introduced to overcome the high variability acoustic index signal. This analysis is in offline mode to discard the outliers of the acoustic index that came from different background noises.

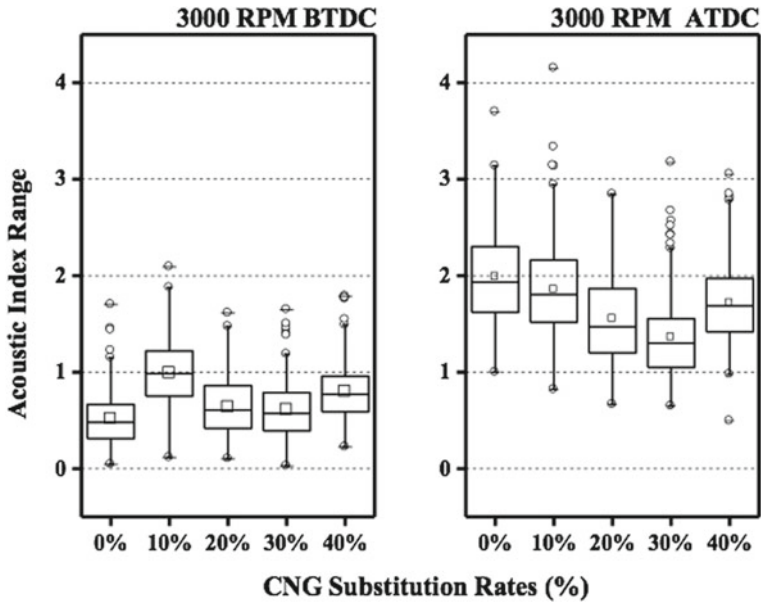


Fig. 34.9 Acoustic index versus CNG substitution rates for 3000 rpm

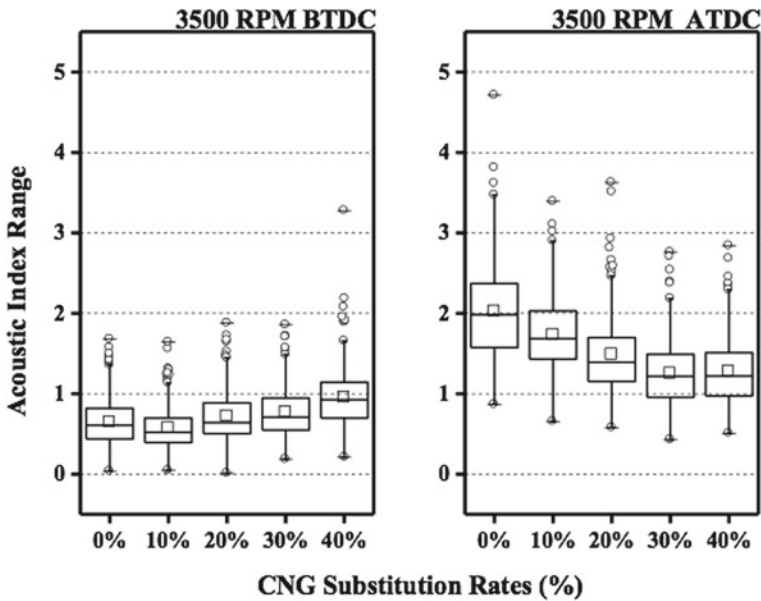


Fig. 34.10 Acoustic index versus CNG substitution rates for 3500 rpm

2. The frequency distribution analysis presents the acoustic index data through its quartile with box plot. Beside to separate the outliers of the acoustic index data, it has provided useful information regard to engine stability.
3. The functionality of knock sensor is enhanced despite it originally used as knock detection.
4. Maximum CNG substitution rate during low engine speeds (<2000 rpm) is 10%. However, higher CNG substitution rates can be achieved for intermediate to high engine speeds.

Acknowledgements The authors wish to thank the Ministry of Higher Education Malaysia (MOHE) and Universiti Tun Hussein Onn Malaysia (UTHM) for partly supporting this research under the Fundamental Research Grant Scheme (FRGS) Vot No.1492 and TIER 1 Grant Vot No. H255.

References

1. Phormula Homepage, <https://phormula.com/product/phormula-ks-3-knock-detector/>. Accessed 2021/09/01
2. PLEX Tuning Homepage, <https://www.plex-tuning.com/products/plex-knock-monitor/>. Accessed 2021/09/01
3. F. Molinaro, F. Castanié, Signal processing pattern classification techniques to improve knock detection in spark ignition engines. *Mech. Syst. Signal Process.* **9**(1), 51–62 (1995)
4. W. Li, F. Gu, A.D. Ball, A.Y.T. Leung, C.E. Phipps, A study of the noise from diesel engines using the independent component analysis. *Mech. Syst. Signal Process.* **15**(6), 1165–1184 (2001)
5. S. Klinchaeam, P. Nivesrangsan, M. Lokitsangthong, Condition monitoring of a small four-stroke petrol engine using vibration signals. *KMITL Sci. Tech. J.* **9**(1), 9–17 (2009)
6. S.B. Devasenapati, V. Sugumaran, K.I. Ramachandran, Misfire identification in a four-stroke four-cylinder petrol engine using decision tree. *Expert Syst. Appl.* **37**(3), 2150–2160 (2010)
7. S. Delvecchio, G. D’Elia, E. Mucchi, G. Dalpiaz, Advanced signal processing tools for the vibratory surveillance of assembly faults in diesel engine cold tests. *J. Vib. Acoust.* **132**(2), 0210081–02100810 (2010)
8. Delvecchio S, D’Elia G, Mucchi E, Di Gregorio R, On the monitoring and diagnosis of assembly faults in diesel engine cold tests: a case study, in *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, ASME Proceedings Series, vol. 1, pp. 3–12, 2010
9. M.J. Mahjoob, A. Zamanian, Vibration signature analysis for engine condition monitoring and diagnosis, in *Proceedings of ISMA2006*, Leuven, Belgium, 2006
10. J.D. Wu, J.C. Chen, Continuous wavelet transform technique for fault signal diagnosis of internal combustion engines. *NDT E Int.* **39**(4), 304–311 (2006)
11. Z. Geng, J. Chen, H.J. Barry, Analysis of engine vibration and design of an applicable diagnosing approach. *Int. J. Mech. Sci.* **45**(8), 1391–1410 (2003)
12. J. Antoni, J. Daniere, F. Guillet, Effective vibration analysis of IC engines using cyclostationarity. Part I-A methodology for condition monitoring. *J. Sound Vib.* **257**(5), 815–837 (2002)
13. J. Antoni, J. Daniere, F. Guillet, R.B. Randall, Effective vibration analysis of IC engines using cyclostationarity. Part II—new results on the reconstruction of the cylinder pressures. *J. Sound Vib.* **257**(5), 839–856 (2002)

14. S. Delvecchio, G. D'Elia, G. Dalpiaz, On the use of cyclostationary indicators in IC engine quality control by cold tests. *Mech. Syst. Signal Process.* **60–61**, 208–228 (2015)
15. R. Mohsin, Z.A. Majid, A.H. Shihnan, N.S. Nasri, Z. Sharer, Effect of biodiesel blends on engine performance and exhaust emission for diesel dual fuel engine. *Energy Convers. Manage.* **88**, 821–828 (2014)
16. M.M. Ismail, F.H. Zulkifli, M. Fawzi, S.A. Osman, Conversion method of a diesel engine to a CNG-diesel dual fuel engine and its financial savings. *ARPN J. Eng. Appl. Sci.* **11**(8), 5078–5083 (2016)
17. Ismail MM, Effect of fuel delivery ratio of diesel-CNG dual fuel engine on performance and emissions. Universiti Tun Hussein Onn Malaysia, 2016
18. A. Iijima, K. Takeda, Y. Yoshida, Z. Lin, H. Shoji, A study of interaction between pressure waves and reaction regions in HCCI combustion accompanied by strong knocking based on high-speed in-cylinder visualization and observation, 26th ICDERS, Boston, MA, USA, pp. 1–6, 2017
19. J.R. Taylor, *An Introduction to Error Analysis the Study of Uncertainties in Physical Measurements*, 2nd edn. (University Science Books, Mill Valley, 1997)
20. P.R. Bevington, D.K. Robinson, *Data Reduction and Error Analysis for the Physical Sciences*, 3rd edn. (McGraw-Hill, NY, 2003)
21. H.W. Coleman, W.G. Steele, *Experimentation, Validation, and Uncertainty Analysis for Engineers*, 4th edn. (Wiley, London, 2018)

Chapter 35

Impulse Breakdown Characteristics of CF₃I–CO₂ at Various Gas Pressure and CF₃I Content



Muhammad Saufi Kamarudin, Noor Mazliza Badrul Sham, A. Haddad, Md Nor Ramdon Baharom, Mohd Fairouz Mohd Yousof, and Nordiana Azlin Othman

Abstract Sulphur hexafluoride (SF₆) is currently the preferred dielectric gas used in medium and high voltage power equipment. However, due to the extremely greenhouse effect by SF₆, the search for alternative gas comprising a fluoro-organic gas with low environmental concern mixed with a buffer gas (N₂, CO₂ and O₂) have received considerable attention from the power industry. Trifluoroiodomethane (CF₃I) is identified to have an excellent insulation performance with a very low GWP can replace SF₆. This paper presents the basic lightning impulse performance of CF₃I mixed with CO₂ as an alternative to SF₆ insulation gas. The 50% breakdown voltage (U₅₀) under both positive and negative polarities were measured at various gas pressure and mixing ratio of CF₃I–CO₂. The results show that U₅₀ of CF₃I–CO₂ gas mixtures increases linearly with the pressure, and the difference increases with gap length. Furthermore, an increment of CF₃I content in the mixtures increases the insulation strength. It was also found that U₅₀ with negative polarity is much higher than that with positive polarity under the needle-plane electrode and vice versa under the plane-plane configuration. The *V-t* characteristics for CF₃I–CO₂ were also analysed under these two different electric fields. The *V-t* characteristics for plane-plane electrode are more distributed along the wavetail of the lightning impulse under both U₅₀ polarities. As opposed to the *V-t* waveforms under needle-plane electrode, the instantaneous breakdowns tend to happen earlier. Overall, the results provide a basis for considering the application of CF₃I–CO₂ as an insulating gas in high voltage apparatus.

M. S. Kamarudin (✉) · N. M. Badrul Sham · M. N. R. Baharom · M. F. Mohd Yousof · N. A. Othman

Department of Electrical Power, Faculty of Electrical and Electronic Engineering, Universiti Tun Hussein Onn Malaysia (UTHM), Parit Raja, 86400 Batu Pahat, Johor, Malaysia
e-mail: saufi@uthm.edu.my

A. Haddad
School of Engineering, Cardiff University, Cardiff, UK

35.1 Introduction

Sulphur hexafluoride (SF_6) is widely recognised as the best insulation gas in high voltage apparatus due to its outstanding insulation strength and arc quenching properties. But, the longtime excessive usage of SF_6 may increase the percentage of its global warming potential (GWP) and have an extremely long atmospheric lifetime [1, 2]. For these reasons, limiting the application of SF_6 , or eventually replaced it by using alternative gases is necessary for the developed countries such as EU, US and China [3].

Much research in recent years has focused on gases or mixed gas, especially the fluorocarbons with halogenated gases or fluorinated compounds, which have similar or better dielectric strength to that of SF_6 . Although some perfluorocarbons (PFCs) and hydrofluorocarbons (HFCs) offer an excellent breakdown strength, their GWP remains high, typically in the range of 5000–12,000 [4, 5]. Their lifetime is also relatively higher than SF_6 , thus limiting them from consideration. One of the most promising low GWP alternatives is trifluoroiodomethane (CF_3I). Its GWP is less than 5, and the insulation strength of CF_3I is 1.2 times higher than that of SF_6 [6, 7].

CF_3I is colorless, odorless and non-flammable have considered being an environmentally friendly insulation gas [8, 9]. Due to the high liquefaction temperature, CF_3I needs to be mixed with buffer gases (N_2 or CO_2) that have relatively low boiling point [10, 11]. Thus, the problem of high liquefaction temperature of CF_3I can be avoided and the mixtures could possess identical features like economical cost, good dielectric performance and provide less environmental concerns.

In this paper, the fundamental lightning impulse withstand of $\text{CF}_3\text{I}-\text{CO}_2$ gas mixtures are measured at a pressure range of 1.0 bar to 2.0 bar. In addition, the effect of various mixture ratios on the insulation performance of $\text{CF}_3\text{I}-\text{CO}_2$ gas mixtures was also analysed. From this study, as the CF_3I content increased, the dielectric strength is more significant in the plane-plane electrode. Meanwhile, in the needle-plane electrode, the increment in the dielectric properties is noticed only in the small gap length.

35.2 Experimental Arrangement and Method

Figure 35.1 shows the complete block diagram of the test setup. The specific details of the lightning impulse generation, structural of pressure vessel, as well as electrode geometry are described in this section.

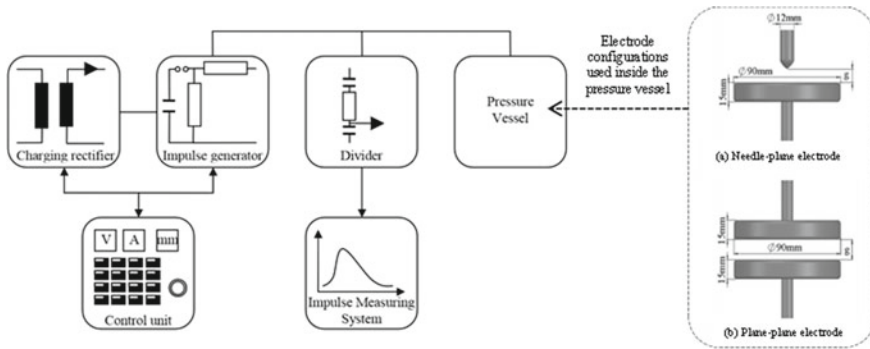


Fig. 35.1 Block diagram of the test setup

35.2.1 Lightning Impulse Test

The impulse voltages were generated by a four (4) stage Marx impulse generator. The maximum output voltage level is 400 kV, delivering 100 kV per stage [12]. The impulse voltage was then measured by a 50 ns rise time capacitive voltage divider and recorded using a digital oscilloscope. The up-and-down method was used to obtain the 50% breakdown voltage (U_{50}) with at least 20 impulse shots for each test condition.

35.2.2 Pressurized Vessel

The cylindrical-shaped pressure vessel was built with a height of 500 mm, having a radius of 250 mm, and a thickness of 10 mm, giving a volume around 0.0982 m³, or 98.2 L. The main constituent material of the vessel is mild steel, a non-alloy structural steel that can withstand specific high pressure and high voltage stresses. Meanwhile, the vessel window is made of polycarbonate, in which able to observe the experimental phenomena easily.

35.2.3 Electrode Configuration

A needle-plane and plane-plane electrode were used in this study. The diameter of the needle electrode is 12 mm, having a tip curvature diameter of 1 mm, while the diameter of the plane electrode is 90 mm, the thickness is 15 mm, and the chamfer radius is 5 mm. All electrodes are made from brass. The electrodes surface was mirror finished to ensure no effect from protrusion, compromising the test results.

35.3 Result and Discussion

This section focuses on both positive and negative lightning impulse breakdown testing. The effect of various gas pressure and different mixture ratios on the impulse breakdown characteristics of $\text{CF}_3\text{I}-\text{CO}_2$ were briefly discussed.

35.3.1 Breakdown Performance of $\text{CF}_3\text{I}-\text{CO}_2$ Mixture Under Different Pressures

Figure 35.2 shows the impulse breakdown voltage-pressure relationship of 30% CF_3I -70% CO_2 gas mixture under positive and negative polarity at different gas pressures and various electrode gap distances. The U_{50} of CF_3I mixture increase with the increasing pressure range between 1.0 bar and 2.0 bar (abs). It is noted that the saturation value is approached more gradually at a higher pressure, especially in 5 cm gap under both U_{50} polarities.

The $V-t$ characteristics in Fig. 35.3 indicate the positive lightning impulse breakdown properties of $\text{CF}_3\text{I}-\text{CO}_2$ gas mixtures at the respective pressure range studied. In the case of positive impulse polarity under a non-uniform (needle-plane electrode) configuration, most of the breakdowns appeared around the wavefront duration corresponding to the peak value, less than 8 μs . A closer examination of the results in Fig. 35.3a reveals that the duration of the breakdown voltage at 1 cm gap length under 1.0 bar pressure occurs near the peak value in the region of 1–2 μs . When the gas pressure is changed to 1.5 bar and 2.0 bar, the breakdown voltages are more dispersed throughout the time scale, reach to 6 μs , as shown in Fig. 35.3b, c, respectively. At a pressure of 2.0 bar, shown in Fig. 35.3c, the $V-t$ plots for a 4 cm gap is almost similar to that for a 5 cm gap with only 2 kV difference. This behaviour shows the saturation phenomenon under positive lightning impulse at higher pressure and gap length.

Figure 35.4 shows the corresponding $V-t$ waveforms of $\text{CF}_3\text{I}-\text{CO}_2$ mixtures under negative polarity in needle-plane electrode configuration. As observed in Fig. 35.2,

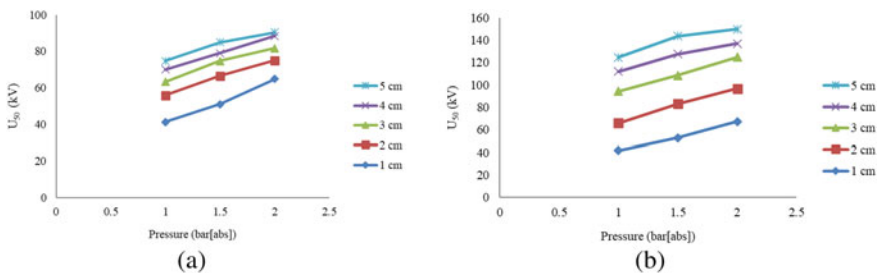


Fig. 35.2 Effects of $\text{CF}_3\text{I}-\text{CO}_2$ pressures on (a) positive, and (b) negative breakdown voltages, U_{50}

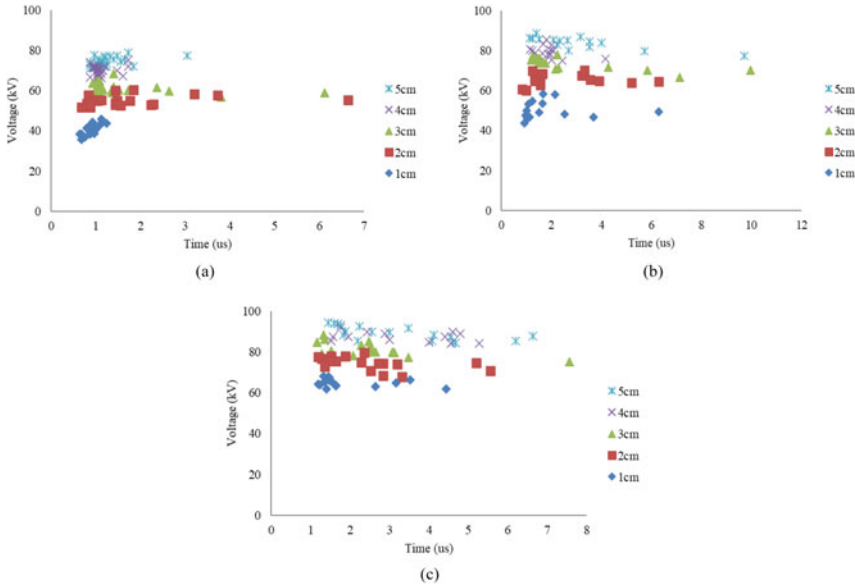


Fig. 35.3 *V-t* characteristics for 30%CF₃I–70%CO₂ under positive impulse at (a) 1.0 bar, (b) 1.5 bar, and (c) 2.0 bar

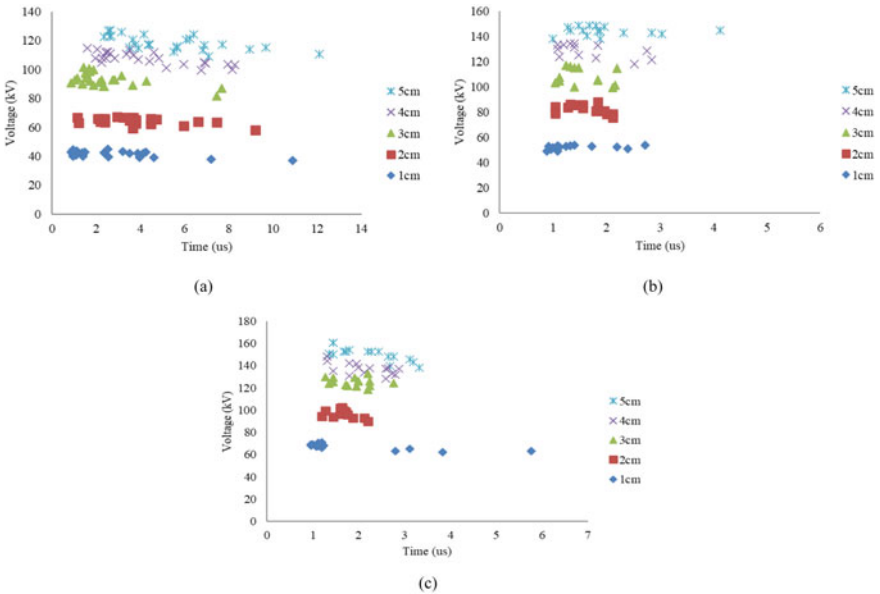


Fig. 35.4 *V-t* characteristics for 30%CF₃I–70%CO₂ under negative impulse at a 1.0 bar, b 1.5 bar, and c 2.0 bar

50% breakdown voltages under negative lightning impulse are significantly higher than those under positive impulses in an expanded range of pressures. The differences in U_{50} increase as the gap increase. Furthermore, the increase rate of U_{50} is more significant over the gap distance between electrodes at all pressures. The increasing trend can be noticed in Fig. 35.4 when compared to Fig. 35.3. A relatively significant time lag was observed at lower pressure of 1.0 bar in the negative impulse under the non-uniform field. However, the $V-t$ waveforms at higher pressure are concentrated near the peak value with a duration of less than 6 μ s.

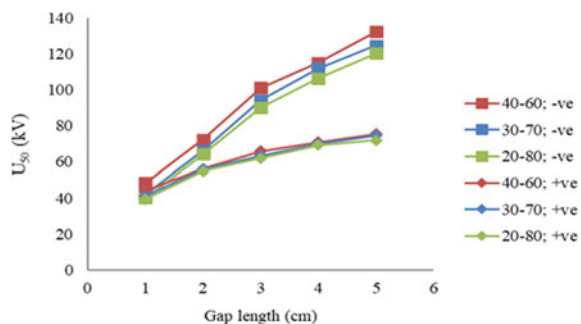
35.3.2 Breakdown Performance of CF_3I-CO_2 Mixture Under Different Pressures

Further impulse breakdown tests were carried out for three (3) different mixture ratios, which are 20%–80%, 30%–70% and 40%–60% of CF_3I-CO_2 at a pressure of 1.0 bar. The tests involved two (2) different electrode arrangements, including the needle-plane and the plane-plane configuration.

Effects of CF_3I Content in Needle-Plane Configuration. The breakdown voltage (U_{50}) at different CF_3I-CO_2 gas mixing ratios for various gap distance is plotted in Fig. 35.5. It is noted that the increment of U_{50} is nearly linear with the increasing electrode gap under positive impulse voltage and gradually increase for negative impulse voltage. However, as expected in a needle-plane electrode, U_{50} under the negative impulse is probable to be higher than those under positive polarity. This phenomenon shows that the electron emission from the cathode surface was remarkable at the higher electric field under a non-uniform configuration. Therefore, U_{50} would have the maximum value to initiate a complete breakdown.

The growth trend of U_{50} curves is obvious under negative impulse polarity with the increase of CF_3I content. The comparison can be made between the highest and the lowest concentration of CF_3I in CF_3I-CO_2 gas mixtures. The U_{50} of 40% $CF_3I-60\%CO_2$ mixture is about 21% higher than that of 20% $CF_3I-80\%CO_2$ at a 1 cm gap distance. For the 3 cm and 5 cm gap, U_{50} of the highest CF_3I content is 12 and

Fig. 35.5 Breakdown voltages at various gap length for different CF_3I-CO_2 mixture ratios in needle-plane electrode



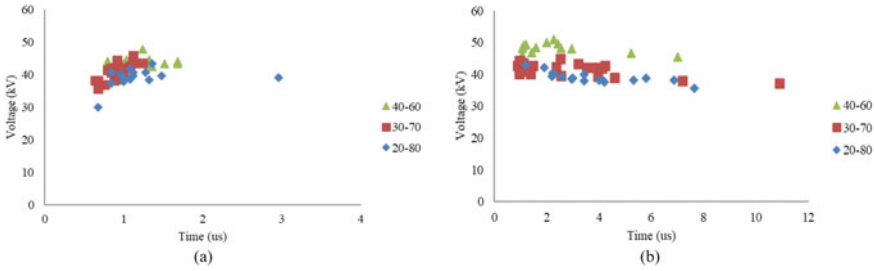


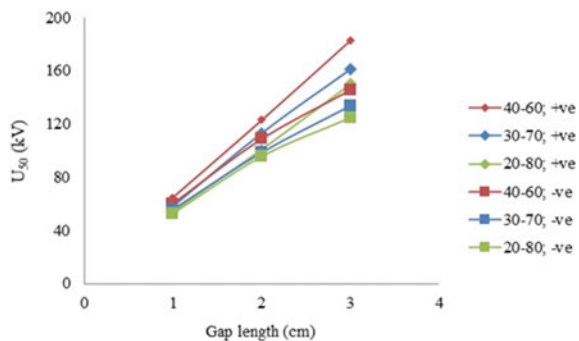
Fig. 35.6 *V-t* characteristics for different CF₃I–CO₂ mixture ratios under **a** positive and **b** negative *U*₅₀ in the needle-plane configuration

10% higher. While a slight effect on the positive impulse breakdown voltages can be seen when changing the proportion of CF₃I electronegative gas from 20 to 40% in CF₃I–CO₂ mixtures. There is only 11% increase of *U*₅₀ at 1 cm gap, 6% at 3 cm gap, and 5% at 5 cm gap distance. From the result, the effect of CF₃I content in the CF₃I–CO₂ mixtures was noted on smaller gaps under both polarities.

Figure 35.6 shows the *V-t* plots of CF₃I–CO₂ under non-uniform field electrode at 1 cm gap for both positive and negative standard lightning impulse. As mentioned in previous discussion, the breakdown voltages under a positive polarity tend to happen earlier, near the peak value in the range of 1–2 μs under 1.0 bar gas pressure. This occurrence can be seen for all CF₃I–CO₂ mixing ratios shown in Fig. 35.6a, caused by the developing speed of discharge under a non-uniform electrode configuration. Under the negative polarity, the duration of the instantaneous voltages is longer and more scattered along the negative *U*₅₀ wavetail shown in Fig. 35.6b. Among the three mixture ratios of CF₃I–CO₂, the *V-t*-plot for 40%CF₃I–60%CO₂ gas mixture is the highest under both impulse polarities.

Effects of CF₃I Content in Plane-Plane Configuration. The effect of CF₃I content in CF₃I–CO₂ mixtures was further observed under a more uniform electric field, which represents by the plane-plane electrode configuration. The tests involved gap lengths between 1 and 3 cm. Based on the data demonstrated in Fig. 35.7, the standard

Fig. 35.7 Breakdown voltages at various gap length for different CF₃I–CO₂ mixture ratios in plane-plane electrode



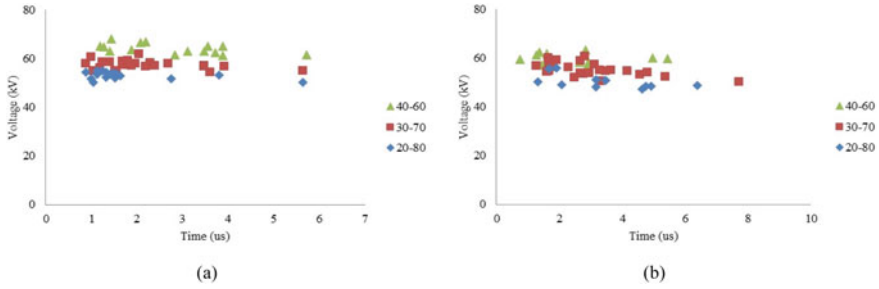


Fig. 35.8 V - t characteristics for different CF_3I - CO_2 mixture ratios under **a** positive and **b** negative U_{50} in the plane-plane configuration

impulse breakdown voltages (U_{50}) under a uniform electrode contrast from those under a non-uniform electrode, which means the U_{50} with positive standard lightning impulse is higher. In all mixtures, U_{50} values increase linearly with increasing the gap distance between electrodes.

Furthermore, it is noted that a slight increment in U_{50} between positive and negative impulse polarity when CF_3I content was increased. Under positive polarity, the relationship between U_{50} and the gap length at different mixing ratios of CF_3I - CO_2 is almost linear. An increase from 20 to 40% of CF_3I in the mixture shows significantly less changes for the slope of the U_{50} . About 20% increase of U_{50} at 1 cm gap, 22% at 2 cm gap, and 22% at 3 cm gap distance. A smaller percentage difference in U_{50} also can be found under negative polarity while varying the proportion of CF_3I content in the mixture. The U_{50} of 40% CF_3I -60% CO_2 mixture is higher than 20% CF_3I -80% CO_2 about 15% at a 1 cm gap, 12% and 10% at a corresponding 3 cm and 5 cm gap.

Figure 35.8 shows the V - t waveforms at 1 cm plane-plane electrode gap for both lightning impulse polarities. In contrast to the non-uniform electrode arrangement, the instantaneous breakdowns in a more uniform field electrode are more distributed along the wavetail time duration under both positive and negative polarity. Among all mixture ratios, the V - t characteristic for CF_3I - CO_2 at 40%-60% remains highest similar to the V - t plots in the needle-plane electrode.

35.4 Conclusion

The dielectric strength of CF_3I - CO_2 mixtures is significant at higher pressures. The 50% breakdown voltages, U_{50} for both lightning impulse polarities increase over the pressure range under the plane-plane and the needle-plane electrode arrangement. When the gap between the electrode is varied, U_{50} of CF_3I - CO_2 mixtures shows the same increment trend as the change in pressure. Furthermore, the increment in CF_3I concentration increase the dielectric strength of CF_3I - CO_2 , obviously in the plane-plane configuration compared to the needle-plane configuration. On the other

hand, the V - t characteristics for CF₃I–CO₂ under uniform electric field (plane-plane electrode) are more dispersed along the U_{50} wavetail duration for both impulse polarities. Unlike the V - t waveforms under the non-uniform field (needle-plane electrode), the instantaneous voltages tend to occur earlier at the peak value of the wavefront duration. However, in all mixture ratios, V - t -characteristic for CF₃I–CO₂ at 40%–60% is the highest under both lightning impulse polarities. Overall, the insulation performance of CF₃I–CO₂ gas mixtures has been proved and can be considered a possible substitute to SF₆ as a new insulating medium.

Acknowledgements The authors gratefully acknowledge the financial supports from the Ministry of Higher Education Malaysia under the Fundamental Research Grant Scheme Vot No. FRGS/1/2018/TK10/UTHM/02/10 and partially sponsored by Universiti Tun Hussein Onn Malaysia.

References

1. S. Xiao, X. Zhang, J. Tang, S. Liu, A review on SF₆ substitute gases and research status of CF₃I gases. *Energy Rep.* **4**, 486–496 (2018)
2. B. Khan et al., Analysis of the dielectric properties of R410A gas as an alternative to SF for high-voltage applications. *High Voltage* **4**(1), 41–48 (2019)
3. M.S. Kamarudin et al., CF₃I gas and its mixtures: potential for electrical insulation, in *International Conference on Large High Voltage Electric Systems (CIGRE) Session 45*, 2014
4. X. Li, H. Zhao, A.B. Murphy, SF₆-alternative gases for application in gas-insulated switchgear. *J. Phys. D: Appl Phys.* **51**(15) (2018)
5. H.S. Kharal, M. Kamran, R. Ullah, M.Z. Saleem, M.J. Alvi, Environment-friendly and efficient gaseous insulator as a potential alternative to SF₆. *Processes* **7**, 1–13 (2019)
6. M.S. Kamarudin, M. Albano, P. Coventry, N. Harid, A. Haddad, A survey on the potential of CF₃I gas as an alternative for SF₆ in high voltage applications, in *45th International Universities Power Engineering Conference (UPEC)*, 2010
7. X. Zhang, S. Xiao, Y. Han, Q. Dai, Analysis of the feasibility of CF₃I/CO₂ used in C-GIS by partial discharge inception voltages in positive half cycle and breakdown voltages. *IEEE Trans. Dielectr. Electr. Insul.* **22**(6), 3234–3243 (2015)
8. X. Li, H. Zhao, J. Wu, S. Jia, Analysis of the insulation characteristics of CF₃I mixtures with CF₄, CO₂, N₂, O₂ and air. *J. Phys. D: Appl. Phys.* **46**(34), 345203 (2013)
9. X. Zhang, S. Xiao, J. Zhou, J. Tang, Experimental analysis of the feasibility of CF₃I/CO₂ substituting SF₆ as insulation medium using needle-plate electrodes. *IEEE Trans. Dielectr. Electr. Insul.* **21**(4), 1895–1900 (2014)
10. S. Zhao, D. Xiao, P. Xue, R. Zhong, Analysis of insulation performance and polar effect of CF₃I/CO₂ mixtures. *IEEE Trans. Dielectr. Electr. Insul.* **25**(4), 1364–1370 (2018)
11. B. Zhou, D. Tan, J. Xue, F. Cai, D. Xiao, Lightning impulse withstand performance of CF₃I–N₂ gas mixture for 252 kV GIL insulation. *IEEE Trans. Dielectr. Electr. Insul.* **26**(4), 1190–1196 (2019)
12. T. Wen et al., Discussion on lightning impulse test waveform according to breakdown characteristics of SF₆ gas gaps. *IEEE Trans. Dielectr. Electr. Insul.* **24**(4), 2306–2313 (2017)

Chapter 36

Optimal Rain Attenuation Prediction Models for Earth-Space Communication at Ku-Band in North Central Nigeria



K. C. Igwe

Abstract Precipitation adversely affects satellite-earth links operating at frequencies ≥ 10 GHz. This paper evaluates five of the best performing rain attenuation models for earth-space links so as to obtain the optimal models for North Central Nigeria. The models considered are ITU-R P.618-9, Bryant, Simple attenuation, Garcia-Lopez and Svjatogor. Also, three elevation angles were considered: 55° (the look angle of most satellite receivers over the Atlantic Ocean Region (AOR) in Nigeria), 23° (the look angle of receivers over the Indian Ocean Region (IOR)) and 42.5° (the look angle of Nigeria's communication satellite (NIGCOMSAT-1R) over the AOR). 2–4 years of 5-min integration time rainfall data was obtained from the Centre for atmospheric research (CAR), Anyigba, Nigeria. The cumulative distribution of rain attenuation at Ku-band was predicted for circularly polarised signals at different percentages of time of the year. Results obtained showed that the ITU-R P.618, Garcia-Lopez and Bryant models performed optimally in this region with attenuation values ranging from 6 to 19 dB at exceedance time percentage of 0.01% in all the stations.

36.1 Introduction

Amongst the debilitating atmospheric effects on satellite-earth links propagating at higher frequency bands (≥ 10 GHz), attenuation by rain is most deleterious to the system's functionality [1–5]. Therefore, it is important to carefully analyse and quantify the extent of attenuation by rain at the frequency of interest [6].

During rainfall, radio waves propagating through the atmosphere at these high frequencies are either absorbed or scattered, thereby resulting to signal reduction [7]. Scattering may in addition cause interference on the radio paths. In the super high frequency (SHF) band, where the wavelength of the radio wave is longer in comparison with the rain drop size, attenuation by absorption will exceed that by

K. C. Igwe (✉)

Department of Physics, Federal University of Technology, Niger State, Minna, Nigeria
e-mail: k.igwe@futminna.edu.ng

scattering. Conversely, in the extreme high frequency (EHF) band and beyond where the wavelength is shorter than the rain drop dimension, attenuation by scattering will be more pronounced [8].

It is difficult to determine the attenuation of electromagnetic signals due to rain because of the unstable nature of rain. Thus, since drop sizes vary temporally and spatially, they do not have a definite arrangement for rainfall rates [9].

Attenuation induced by rain is an important propagation effect that should not be neglected in satellite system design. For optimal analysis of rain attenuation, it is imperative that the corresponding rain rate is accurately evaluated [10]. Some recent works on rain rate attest to this deduction [11–14]. Satellite beacon signals and radiometers are often used to measure rain-induced attenuation but such measurements are done only in very few locations of the world and thus impossible to directly apply the obtained results to all locations. Consequently, in order to make inputs for system margin calculations to be available in every region of the world, meteorological data have been obtained and different rain attenuation models have been developed [15].

In Nigeria, NIGCOMSAT-1R operates on Ku and Ka bands. It is therefore important to know the extent of rain-induced attenuation on satellite-earth links in various locations so as to guide satellite system designers on enhancing the quality of local network contents [16].

This paper therefore evaluates the degree of attenuation induced by rain on satellite-earth communication and further investigates the optimal rain attenuation prediction models in North Central Nigeria by using rainfall data of 5-min integration time as against daily rainfall data used in [16].

36.2 Background

36.2.1 Rainfall Rate Model

Lavergnat-Gole model. The application of this model is given in [17]. It converts cumulative distribution of rain rate from integration time t_1 , to the desired integration time t_2 . This is done by using a conversion factor given as the ratio of the integration times:

$$P_2(R_2) = CF^a P_1(R_1) \quad (36.1)$$

where

$$CF = \frac{t_2}{t_1} \quad (36.2)$$

and

$$R_2 = R_1 / C F^a \quad (36.3)$$

P_1 and P_2 are the cumulative probabilities obtained with rain gauges of t_1 (min) and t_2 (min) integration times respectively, while R_1 (mm/h) and R_2 (mm/h) are the rain rates for P_1 and P_2 respectively.

Parameter 'a' depends on the region of interest. For the temperate region, it is 0.115 [18], while a value of 0.143 was deduced by [19] for the tropical climatic region. This model was used for the rain rate computation since it has recently been adjudged the optimal rain rate model for the North Central region of Nigeria [20].

36.2.2 Rain Attenuation Models

For the rain attenuation models, five amongst the best performing models were selected for rain attenuation estimation in this work. These are the globally accepted ITU-R P.618-9 [21], Bryant [22], Simple attenuation [23], Garcia-Lopez [24] and Svjatogor models [25]. The ITU-R P.618-9 model is explained in detail here, while the other four models are only defined because of space constraint. Their detailed explanation can be found in [16].

ITU-R P. 618-9 model. Rain rate at 0.01% exceedance is used for rain attenuation prediction in this model. Then an adjustment factor is applied for prediction at other percentage exceedances. The parameters required for the procedures are:

$R_{0.01}$ (mm/h)	point rainfall rate at 0.01% of a year.
h_S (km)	Station height above mean sea level
θ (degrees)	Elevation angle
φ : (degrees)	Latitude of the station
f (GHz)	Frequency

The following steps are required for the computation:

Step 1 Compute the rain height, h_R from ITU-R P. 839-3 [26]:

$$h_R = h_o + 0.36km \quad (36.4)$$

where h_o is 0 °C isotherm height above mean sea level.

Step 2 Determine the slant path length, L_S from:

$$L_S = \frac{h_R - h_S}{\text{Sin}\theta} \quad (36.5)$$

Step 3 Compute the horizontal projection, L_G of slant path length using:

$$L_G = L_S \cos \theta \tag{36.6}$$

Step 4 Input $R_{0.01}$ (mm/h).

Step 5 Calculate the specific attenuation, $\gamma_{R0.01}$ (dB/km) at 0.01% from:

$$\gamma_{R0.01} = k R_{0.01}^\alpha \tag{36.7}$$

where k and α are frequency functions determined from ITU-R P.838-3 [27].

Step 6 Compute the horizontal reduction factor, $r_{h0.01}$ using:

$$r_{h0.01} = \frac{1}{1 + 0.78 \sqrt{\left(\frac{L_G \gamma_{R0.01}}{f}\right) - 0.38[1 - \exp(-2L_G)]}} \tag{36.8}$$

Step 7 Calculate the vertical adjustment factor, $v_{0.01}$ (km):

$$L_R = \frac{L_G r_{0.01}}{\cos \theta}, \text{ for } \rho > \theta \tag{36.9}$$

otherwise,

$$L_R = \frac{H_R - H_S}{\sin \theta}, \text{ for } \rho \leq \theta \tag{36.10}$$

where

$$\rho = \tan^{-1} \left(\frac{H_R - H_S}{L_G r_{h0.01}} \right) \tag{36.11}$$

therefore,

$$v_{0.01} = \frac{1}{1 + \sqrt{\sin \theta} \left[31 \left(1 - \exp\left(-\frac{\theta}{[1+\sigma]}\right) \right) \right] \frac{\sqrt{L_G \gamma_{R0.01}}}{f^2} - 0.45]} \tag{36.12}$$

where $\sigma = 36 - |\varphi|$, for $|\varphi| < 36^\circ$; for $|\varphi| \geq 36^\circ$, $\sigma = 0$

Step 8 The effective path length L_{eff} (km) is calculated from:

$$L_E = L_R v_{0.01} \tag{36.13}$$

Step 9 Compute the predicted rain attenuation at 0.01% of an average year using:

$$A_{0.01} = \gamma_{R0.01} L_E \tag{36.14}$$

Step 10 Obtain the attenuation for other percentage exceedances from:

$$A_p(dB) = A_{0.01} \left(\frac{p}{0.01} \right)^{-[0.655+0.033\ln(p)-0.045\ln(A_{0.01})-z\sin\theta(1-p)]} \quad (36.15)$$

where p is the percentage probability of interest. z is obtained from:

$$z = 0 \text{ if } p \geq 1\% \text{ or } |\varphi| \geq 36^\circ; \quad (36.16)$$

$$z = -0.005(|\varphi| - 36) \text{ if } p < 1\% \text{ and } |\varphi| < 36^\circ \text{ for } \theta \geq 25^\circ; \quad (36.17)$$

else

$$z = -0.005(|\varphi| - 36) + 1.8 - 4.25\sin\theta, \text{ for } \theta < 25^\circ \text{ and } |\varphi| < 36^\circ \quad (36.18)$$

Bryant Model. The Bryant attenuation model uses the ‘effective rain cell’ and ‘variable rain height’ concept to compute rain attenuation.

Simple Attenuation Model. This model also uses point rainfall rate for prediction. As an improvement from an earlier version, the effect of wave polarisation is now incorporated [28]

Garcia-Lopez Model. The Garcia-Lopez attenuation model is an extension of the one proposed for terrestrial links. In this model, separate coefficient values are adopted for tropical regions during computation.

Svjatogor Model. This is a type of rain attenuation model whose effective rain height is a function of the rain intensity.

36.3 Methodology

Rainfall at 5-min integration time was collected at the Tropospheric data acquisition network (TRODAN) located at the mini campus of Federal University of Technology, Minna, Nigeria. This data acquisition station was set up by the Centre for atmospheric research (CAR), Anyigba, Nigeria. Data from four other Universities in the North Central region where the TRODAN weather stations are also situated were used in this work. These are Kogi State University, Anyigba, Benue State University, Makurdi and the Universities of Abuja and Jos. The rainfall data was measured by the Campbell CR-1000 data logger. Figure 36.1 shows the TRODAN weather station.

The Lavergnat and Gole model outlined in Eqs. (36.1)–(36.3) was used to estimate the rainfall rate (mm/h), while the ITU-R P.618-9 outlined in Eqs. (36.4)–(36.18), the Bryant, Garcia-Lopez, Simple attenuation and Svjatogor models were used to predict the rain attenuation. Circular polarisation was considered and the Ku-band downlink centre frequency of 12.675 GHz was used in the analysis. Also, three elevation angles were considered. These are 55° (Look angle of satellite receivers



Fig. 36.1 **a** The TRODAN weather station (Campbell data logger and solar panel). **b** Tipping bucket rain gauge

over the AOR in Nigeria), 42.5° (Look angle of NIGCOMSAT-1R over the AOR) and 23° (Look angle of receivers over the IOR).

Table 36.1 lists the input parameters needed for the rain attenuation prediction models.

Table 36.1 Input parameters for the rain attenuation prediction models

Rain model	λ	H_S	θ	f	k, α	$R_p(p)$	$R_{0.01}$
Bryant		✓	✓		✓	✓	
Garcia-Lopez	✓	✓	✓		✓	✓	
ITU-R P.618-9	✓	✓	✓	✓	✓		✓
SAM	✓	✓	✓		✓	✓	
Svjatogor		✓	✓		✓	✓	

where λ : latitude of the station ($^\circ$), H_S : altitude of the station (km), θ : elevation angle ($^\circ$), f: frequency (GHz), k and α : frequency and polarisation dependent coefficients [27], p: time percentage of the year (%), $R_p(p)$: point rainfall rate distribution (mm/h), $R_{0.01}$: point rainfall rate at 0.01% (mm/h)

Table 36.2 Computed rainfall rate at 0.01%

Station	Lat (°N)	Lon (°E)	Altitude (m)	$R_{0.01}$ (mm/h)
Minna	9.54	6.54	249	75.50
Anyigba	7.25	7.18	420	37.80
Makurdi	7.70	8.50	142	75.50
Abuja	9.00	7.28	334	44.00
Jos	9.58	8.57	1110	75.50

36.4 Results and Discussion

36.4.1 Rainfall Rate Estimation

The station characteristics, including the computed rainfall rate $R_{0.01}$ for the North Central region are given in Table 36.2.

36.4.2 Rain Attenuation Prediction

The five predefined models were used to predict rain attenuation at the five stations and comparison of the results were made based on the ITU-R P.618 model. This is because the ITU-R P.618 model has been widely accepted as a very accurate method for estimating rain attenuation on satellite-earth link systems all over the world. When measured data are not available, developed models are usually compared with it [29] since it has been shown that the ITU-R P.618 prediction model agrees closely with measured attenuation values [30–33].

Tables 36.3, 36.4 and 36.5 give values of the relevant input parameters computed using ITU-R P.618-9 model for the three elevation angles.

Figure 36.2a–e show the cumulative distribution of predicted rain attenuation for the stations at 55° elevation angle

From Fig. 36.2, it can be seen that predictions by Bryant and Garcia-Lopez models are close to that of the ITU-R P.618-9 model, though with little deviations of about

Table 36.3 Computed input parameters at 55° elevation angle

Station	h_R (km)	L_S (km)	L_G (km)	L_E (km)
Minna	4.79	5.54	3.18	3.61
Anyigba	4.75	5.28	3.03	5.04
Makurdi	4.76	5.64	3.23	3.68
Abuja	4.76	5.40	3.10	4.72
Jos	4.76	4.45	2.55	3.22

Table 36.4 Computed input parameters at 42.5° elevation angle

Station	h_R (km)	L_S (km)	L_G (km)	L_E (km)
Minna	4.79	6.72	4.95	3.67
Anyigba	4.75	6.41	4.72	5.25
Makurdi	4.76	6.84	5.04	3.75
Abuja	4.76	6.55	4.83	4.89
Jos	4.76	5.40	3.98	3.30

Table 36.5 Computed input parameters at 23° elevation angle

Station	h_R (km)	L_S (km)	L_G (km)	L_E (km)
Minna	4.79	11.61	10.69	5.02
Anyigba	4.75	11.08	10.20	7.22
Makurdi	4.76	11.82	10.88	5.14
Abuja	4.76	11.33	10.43	6.70
Jos	4.76	9.33	8.59	4.51

2–5 dB. It can also be observed that stations that recorded the same rainfall rates at 0.01% have close attenuation values, while the attenuation values are different in stations with different point rainfall rates. This can be noticed in Minna and Makurdi that both recorded 75.5 mm/h at 0.01%. The computed attenuation by the ITU-R P.618-9 model is 13.5 dB for Minna and 13.8 dB for Makurdi. Jos station, with the same value of rain rate also recorded very close attenuation value of 12.0 dB. This trend is replicated by the Bryant and Garcia-Lopez models. Stations like Abuja and Anyigba that recorded lower point rainfall rates had lower rain attenuation computed by the models. The Simple attenuation and Svjatogor models underestimated the rain attenuation as predicted values deviated greatly from those of the other models. The attenuation values computed by the best performing models are further presented in Table 36.6 for clearer illustration.

It is clearly shown in Table 36.6 that these three models also predicted closely at the other time percentage exceedances of 0.001, 0.1 and 1%.

The cumulative distribution of predicted rain attenuation at 42.5° elevation angle are presented in Fig. 36.3a–e.

From Fig. 36.3, it is observed that predicted rain attenuation values are a bit higher at 42.5° elevation angle when compared with values obtained at 55°. The deviation in values between the two elevation angles is about 2–5 dB for the entire distribution. Again, values predicted by the Bryant and Garcia-Lopez models are close to the ITU-R P.618, while computed values from SAM and Svjatogor models are low.

Figure 36.4a–e show the cumulative distribution of predicted rain attenuation at 23° elevation angle

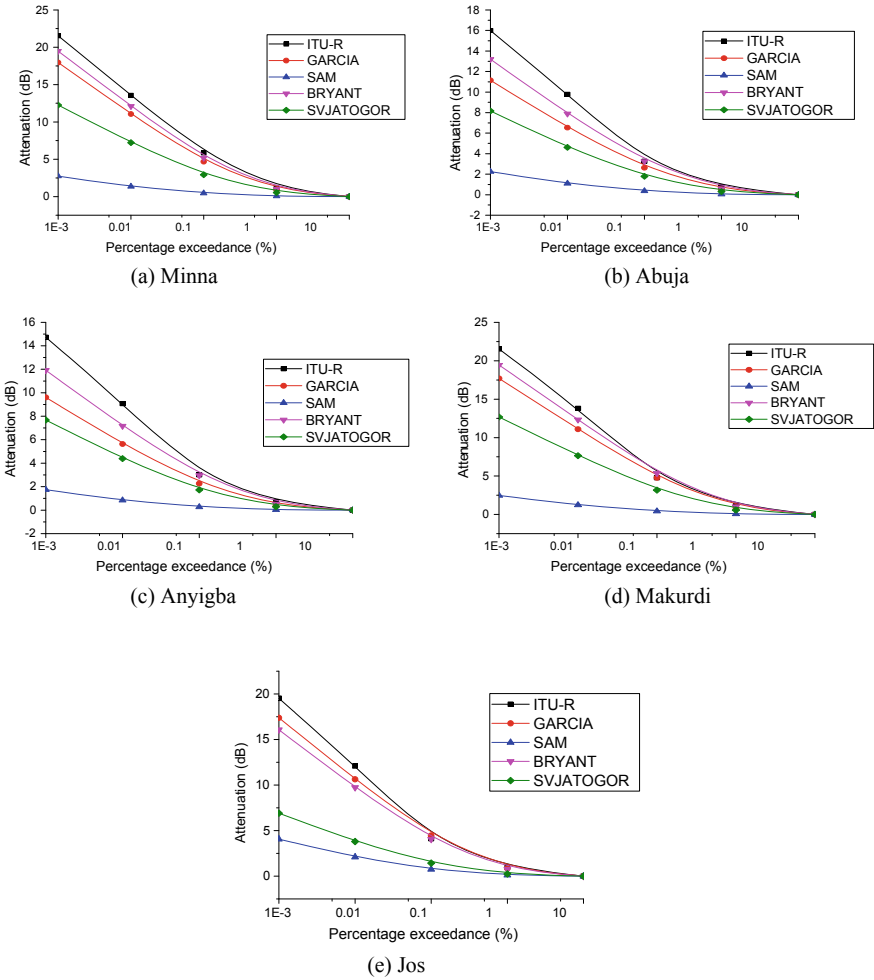


Fig. 36.2 Predicted attenuation at 55° elevation angle

From Fig. 36.4, It is observed that there is an increase in the computed attenuation values at 23° elevation angle. These higher values are due to the longer path length of the rain region associated with this lower angle. This implies that satellite-earth links propagating at this elevation angle will suffer more signal outage than those propagating at higher elevation angles.

Table 36.6 Predicted rain attenuation for the three best attenuation models

% Exceedance	Stations	ITU-R P.618-9 (dB)	Bryant (dB)	Garcia-Lopez (dB)
0.001	Minna	21.5	19.5	18.0
	Abuja	16.0	13.0	11.0
	Makurdi	21.5	19.5	18.0
	Anyigba	15.0	12.0	9.0
	Jos	19.5	16.0	17.0
0.01	Minna	13.5	12.0	11.0
	Abuja	10.0	8.0	6.5
	Makurdi	13.8	12.0	11.0
	Anyigba	9.0	7.0	5.6
	Jos	12.0	9.7	10.6
0.1	Minna	6.0	5.0	5.0
	Abuja	3.3	3.2	2.6
	Makurdi	5.0	5.0	5.0
	Anyigba	3.0	3.0	2.3
	Jos	4.0	4.0	4.5
1	Minna	1.1	1.0	0.9
	Abuja	0.8	0.6	0.5
	Makurdi	1.2	1.0	0.9
	Anyigba	0.7	0.5	0.4
	Jos	1.0	0.8	0.8

36.5 Conclusion

Rain attenuation predictions on satellite-earth links at Ku-band middle frequency operating at 55°, 42.5° and 23° elevation angles were made using five different rain attenuation models. For the computed rain attenuation, it has been shown that higher attenuation was experienced at lower percentage of time exceedance, while lower attenuation was experienced at higher percentage of time exceedance. Also, for the three elevation angles considered, rain attenuation was highest at 23°, followed by 42.5°, while 55° elevation angle had the lowest computed attenuation (though with slightly lower values than at 42.5°). This implies that higher attenuation is experienced at lower elevation angles, while lower attenuation is experienced at higher elevation angles. In all the stations, attenuation generally ranged from 6 to 14 dB at 55° elevation angle, 6 to 14 dB at 42.5° elevation angle and from 6 to 19 dB at 23° elevation angle at time percentage exceedance of 0.01% for models with close predictions.

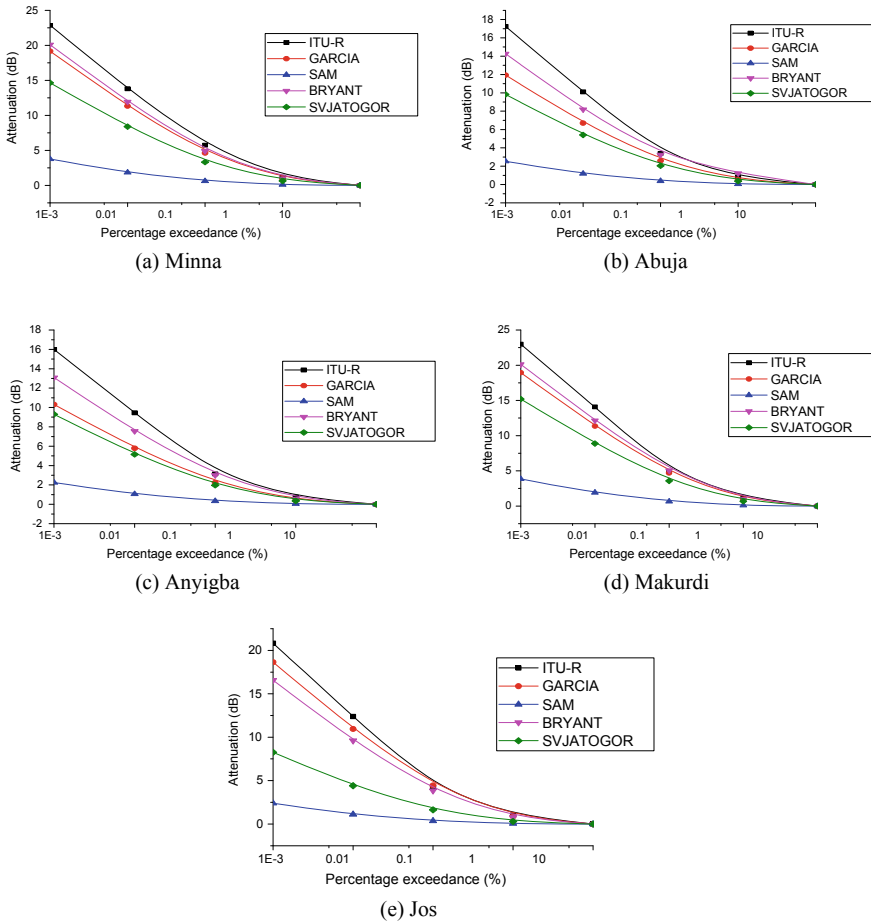


Fig. 36.3 Predicted attenuation at 42.5° elevation angle

In addition, results obtained have shown that three of the five models, the ITU-R P.618, Garcia-Lopez and Bryant were in good agreement, while predictions from the Simple attenuation and Svjatogor models underestimated the predicted rain attenuation values. Thus, the ITU-R P.618, Garcia-Lopez and Bryant models can be used to satisfactorily predict rain attenuation in North Central Nigeria.

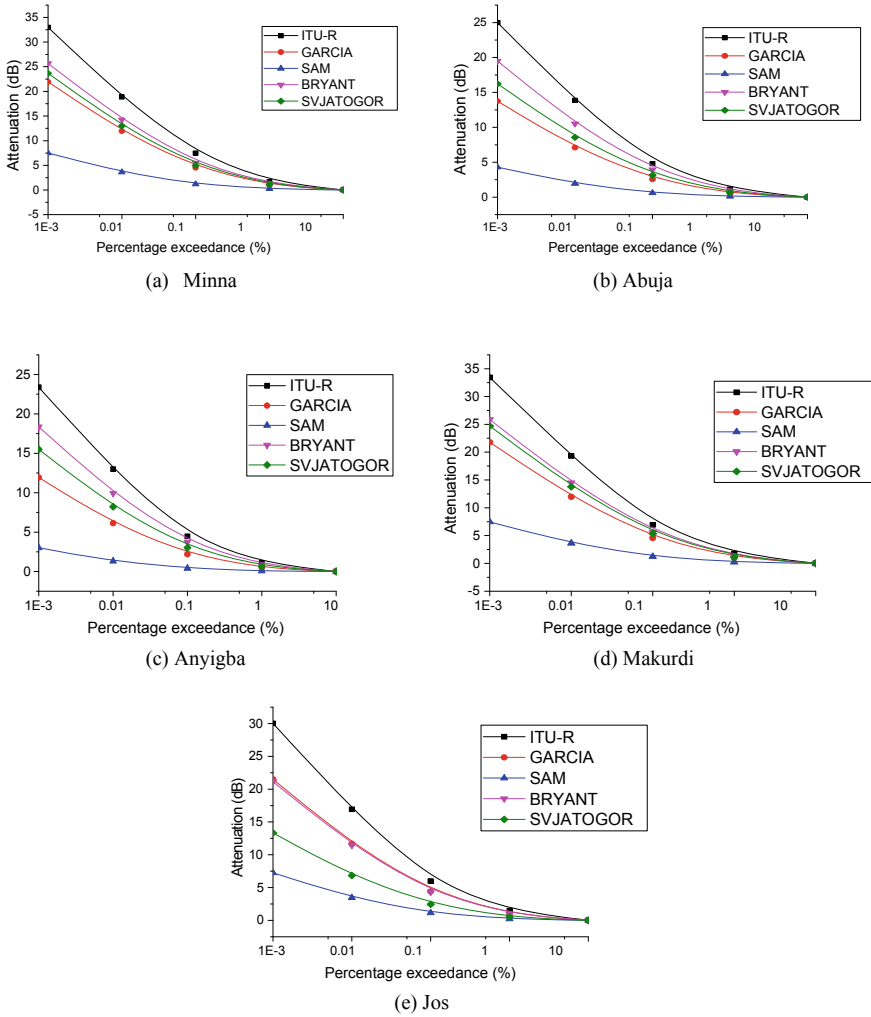


Fig. 36.4 Predicted attenuation at 23° elevation angle

Acknowledgements The author appreciates the Centre for atmospheric research, Anyigba, Nigeria for providing the rainfall measuring equipment used in this research.

References

1. R.K. Crane, Rain attenuation models: Attenuation by clouds and rain, in *Propagation handbook for wireless communication system design*. (CRC Press, USA, 2003), pp. 225–280
2. J.S. Mandeep, Comparison of rain rate models for equatorial climate in South East Asia. *Geofizika* **28**, 265–274 (2011)
3. S. Shrestha, J.J. Park, D.Y. Choi, Rain rate modeling of 1 min from various integration times in South Korea. *SpringerPlus* **5**(1), 1–34 (2016)
4. S. Shrestha, D.Y. Choi, Characterization of rain specific attenuation and frequency scaling method for satellite communication in South Korea. *Int. J. Antennas Propag.* **2017**, 1–16 (2017)
5. M.S. Hossain, M.A. Islam, Estimation of rain attenuation at EHF bands for earth-to-satellite links in Bangladesh. *International Conference on Electrical, Computer and Communication Engineering*, Cox's Bazar, Bangladesh, pp 589–593 (2017)
6. S. Shrestha, D.Y. Choi, Diurnal and monthly variations of rain rate and rain attenuation on Ka-band satellite communication in South Korea. *Prog. Electromag. Res.* **B 80**, 151–171 (2018)
7. L. Rong, A new method of uplink power compensation of rain attenuation of satellite communication system. *International Conference on Automation, Mechanical Control and Computational Engineering*, pp. 2181–2185 (2015)
8. M.P.M. Hall, *Effects of the troposphere on radio communication* (Peter Peregrinus Limited, U.K., 1979)
9. M. Tamosiunaite, M. Tamosiuniene, A. Gruodis, S. Tamosiunas, Prediction of electromagnetic wave attenuation due to water in the atmosphere. 1. Attenuation due to rain. *Innovative Infotechnologies Sci. Bus. Educ.* **2**(9), 3–10 (2010)
10. R. Oktaviani, Marzuki, Estimation of rainfall rate cumulative distribution in Indonesia using global satellite mapping of precipitation data. *International Conference on Basic Science and its Application*, pp. 259–265 (2019)
11. Y. Ng, M. Singh, J. Singh, V. Thiruchelvam, Performance analysis of 60-min to 1-min integration time rain rate conversion models in Malaysia. *J Atm and Solar-Terr Phys* **167**, 13–22 (2017)
12. S. Shrestha, D. Choi, Study of 1-min rain rate integration statistics in South Korea. *J Atm Solar-Terr. Phys.* (2017)
13. I. Rafiqul, M. Alam, A.K. Lwas, S.Y. Mohamad, Rain rate distributions for microwave link design based on long term measurement in Malaysia. *Indonesia J. Elect. Eng. Comp. Sci.* **10**(3), 1023–1029 (2018)
14. R. Singh, R. Acharya, Development of a new global model for estimating one-minute rainfall rate. *IEEE Trans Geosci Remote Sens* **56**(11), 6462–6468 (2019)
15. COST 225, Radiowave propagation modelling for SatCom services at Ku-band and above. Final report, European Space Agency, the Netherlands, pp. 74–97 (2002)
16. K.C. Igwe, O.D. Oyedum, M.O. Ajewole, A.M. Aibinu, Evaluation of some rain attenuation prediction models for satellite communication at Ku and Ka bands. *J. Atm. Solar-Terr Phys.* **188**, 52–61 (2019)
17. J. Lavergnat, P.A. Gole, Stochastic raindrop time distribution model. *J Appl Met* **37**, 805–818 (1998)
18. C. Ito, Y. Hosoya, *Proceedings of ISAP'04, Sendai, JAPAN*, pp. 1361–1364 (2004)
19. L.D. Emiliani, L. Luini, C. Capsoni, Analysis and parameterization of methodologies for the conversion of rain-rate cumulative distributions from various integration times to one minute. *IEEE Ant. Propag. Mag.* **51**(3), 70–80 (2009)
20. K.C. Igwe, O.D. Oyedum, M.O. Ajewole, A.M. Aibinu, J.A. Ezenwora, Performance evaluation of some rain rate conversion models for microwave propagation studies. *Adv. Space Res.* **67**, 3098–3105 (2021)
21. ITU-R, *Propagation data and prediction methods required for the design of earth-space telecommunication systems. Rec P 618-9* (International Telecommunication Union, Geneva, 2007)

22. G.H. Bryant, I. Adimula, C. Riva, G. Brussard, Rain attenuation statistics from rain column, diameters and heights. *Int. J. Sat. Commun.* **19**, 263–283 (1999)
23. W.L. Stutzman, W.K. Dishman, Correction to a simple model for the estimation of rain-induced attenuation along earth-space paths at millimeter wavelengths. *Rad. Sci.* **19**, 946 (1984)
24. J.A. Garcia-Lopez, J.M. Hernando, J. Selga, Simple rain attenuation method for satellite radio links. Year-to-year variability of rainfall for microwave applications in the USA. *IEEE Trans Ant Propag* **36**(3), 444–448 (1988)
25. L. Svjatogor, Prostranstvennaia korelacia vypadenjija dozdznej vdol zemnoj poverchnostji (in Russian), Symposium expertov stran uchastnic programmy. INTERKOSMOS (Interkosmos symposium, theme 5 of the established telecommunication working group, Dresden, GDR, 1985)
26. ITU-R, *Rain height model for prediction methods. Rec P 839-3* (International Telecommunication Union, Geneva, 2001)
27. ITU-R, *Specific attenuation model for rain for use in prediction methods. Rec P 838-3* (International Telecommunication Union, Geneva, 2005)
28. G. Immadi, S.K. Kotamraju, M.V. Narayana, H. Khan, G. Viswanath, I. Avinash, Measurement of rain attenuation for Ku band satellite signal in tropical environment using DAH, SAM models. *ARNP J. Eng. Appl. Sci.* **10**(4), 1717–1722 (2015)
29. Y.I.O. Abayomi, N.H.H. Khamis, Rain attenuation modelling and mitigation in the tropics: brief review. *Int. J. Elect. Comp. Eng.* **2**(6), 748–757 (2012)
30. Y.S. Choi, J.H. Lee, J.M. Kim, Rain attenuation measurements of the KoreaSat beacon signal on 12 GHz. CLIMPARA '98, Ottawa, Canada, pp. 208–211 (1997)
31. L.D. Emiliani, E. Agudelo, E. Gutierrez, J. Restrepo, C. Fradique-Mendez, Development of rain-attenuation and rain-rate maps for satellite system design in the Ku- and Ka-bands in Columbia *IEEE Ant. Propag. Mag.* **46**(6), 54–68 (2004)
32. J.S. Mandeep, J.E. Allnut, Rain attenuation predictions at Ku-band in South East Asia countries. *Prog. Electromag. Res. PIERS* **76**, 65–74 (2007)
33. P. Panchal, R. Joshi, Performance analysis and simulation of rain attenuation models at 12–40 GHz band for an earth space path over Indian cities. *7th International conference on Communication Computer Virtualization*, 79, pp. 801–808 (2016)

Chapter 37

Investigation on Water Losses in Reticulation System at UTHM Campus



Sumitha Selami and Sabariah Musa

Abstract Water loss in water distribution systems is a widespread and serious issue with environmental and economic impacts. There have been reports of high water rates on campus. The objective of the study is to use the water balance method to compute the percentage of water losses and to assess the effectiveness of software created by the American Water Works Association (AWWA) to detect water losses. Throughout the experiment, which took place across the UTHM campus area, the volume of water delivered and the percentage of losses were recorded. This study will look into the function of a bulk and volumetric meter that was previously installed in the water delivery system. Water loss is calculated throughout a seven-week assessment period as the difference between the system's input volume and all customers' billed consumption, and American Water Works Association (AWWA) software is used to review the current yearly water consumption value and cost each month. The proportion of losses as well as the value for Area 1 is 20.85% for 7 weeks, which is less than 25% by international standards, according to the findings of this study. The data also revealed significant problems in Area 2, where water loss exceeded 25% (218.11% for 7 weeks), despite the fact that the causes were mistaken water meters, damaged meters, restricted regions, and other damaged characteristics, not water loss. Aside from that, the AWWA programme calculates the percentage of losses and assesses the American Water Works Association (AWWA) platform's effectiveness in the investigation. The software calculated the annual percentage loss and found that it will be 5.5% in 2020, which is less than 25%. Due to the low percentage of losses, it is assumed that there is no water shortage on campus, and the water balancing method can be altered and combined with the water signal and Acoustic test.

S. Selami · S. Musa (✉)

Department of Civil Engineering, Faculty of Civil Engineering and Built Environment, Universiti Tun Hussein Onn Malaysia, 86400 Batu Pahat, Malaysia
e-mail: sabariah@uthm.edu.my

S. Musa

Faculty of Civil Engineering and Built Environment, Eco-Hytech Research Centre, Universiti Tun Hussein Onn Malaysia, 86400 Batu Pahat, Malaysia

37.1 Introduction

37.1.1 Water Leakage

Globally, the demand for water is growing, and supplies are decreasing. Water access is key to life, stability, and all human activities [1]. Water leakage from the pipe network has always been a *bête noire* of the operating engineer, a characteristic of operational management for a long time, including in countries with well-developed networks and best operating practices [2]. Water losses may occur at certain points of the water distribution system (WDS) including septic tanks, power stations, water meters, service reservoirs and billing systems. Water losses are typically the main part of the so-called Non-Revenue Water (NRW), which is water that does not produce revenue for the water utility.

A system-wide water audit, known as the water balance worldwide, has provided quantitative measures of the overall volume of water consumed. The audits provide a useful examination of the various components of usage and losses that are used to determine the service's efficiency, funding, and management of its water supply. In addition, water audits are necessary for the planning of additional leak management operations [3].

The investigation's objective is to calculate the percentage of water losses using the water balancing method and to assess the use of the American Water Works Association (AWWA) software for detecting water losses. Universiti Tun Hussein Onn Malaysia (UTHM) water delivery system was used for this investigation. The investigation used bulk and volumetric meter to find the volume of water supplied in the area. This study also provided recommendation to detect the leaks and make any necessary fixes before the problem gets worse.

37.1.2 Water Balance Method

The apparent losses (AL) components are measured in this procedure, and the quantity of real losses (RL) is calculated by the number of water leaks. The amount of NRW can be determined using the unit's input volume and billed consumption (BC). After subtracting the volume of non-revenue water from the unbilled authorized water usage (UAC), the quantity of water loss (WL) is calculated. The apparent losses (AL) are then calculated and/or assumed by determining and/or assuming their sub-components, namely meter inaccuracies, error management data, and illegal usage of unbilled consumption (UC). Water meter inaccuracies can be calculated on the basis of meter readings at different flow rates, reflecting typical consumer water use and instruction manuals [4].

Smart Water Meters

Water meters must be manually interpreted over a set period of time, which can be a time-consuming and labor-intensive operation. For most Australian households, meter readings are taken on a quarterly and annual basis. Although this may suffice for billing reasons, there is no information about human water use habits, pollution, or seasonal fluctuations [5]. Water supply management needs accurate, timely, and productive evaluation and control of water consumption activities [6]. Many countries, including Australia, have recently adopted the Smart Water Metering (SWM) technology [7]. Water systems can use SWM technology to read water meters centrally and at varied intensities, as well as in a format that can be used for a variety of reasons, such as output and usage control, leakage prevention, and water retention [7, 8]. The ability to record, analyse, and distribute water usage statistics to water consumers in near real time has the potential to cause significant changes in water usage trends [5]. Smart water meters are frequently utilised to increase water consumption input [9]. For more than two decades, Malaysia's federal and state governments have made water management a primary concern.

Air Selangor (AIS) is in charge of water purification and distribution for the state of Selangor and its bordering territories of Kuala Lumpur, Putrajaya, and Cyberjaya, which total 8.4 million inhabitants. On the Smart Water Integrated Management System (SWIMS), Air Selangor (AIS) collaborated with Telekom Malaysia Research and Development (TM R&D), SWIMS is a network technology agnostic digital solution that allows Air Selangor AIS tools to read and store any customer meter data effectively, allowing them to get the most out of it.

American Water Works Association (AWWA) Softwarey

The AWWA established the Water Audit Program, which was used in the current study, based on the given water balance of the International Water Association (IWA) [10]. The software offers a fantastic Microsoft Office suite with a wide range of worksheets. The application calculates spending, income, and NRW, as well as a complete water balance, once all of the data is entered. The data was entered into the software, and the results were generated and shown in an international-style water balance sheet. The water balance values were calculated using measurements and evaluations of the water input unit, water intake, and water depletion.

37.2 Methods

This study presents a conceptual model and methods for determining the proportion of water losses using a water balance form, as well as evaluating the efficiency of the model using AWWA software for water loss investigation. For this purpose, the water reticulation map is used to classify the water distribution system and to collect the total water generated and the total water consumed in Universiti Tun Hussein Onn Malaysia for each region using a bulk and volumetric meter.

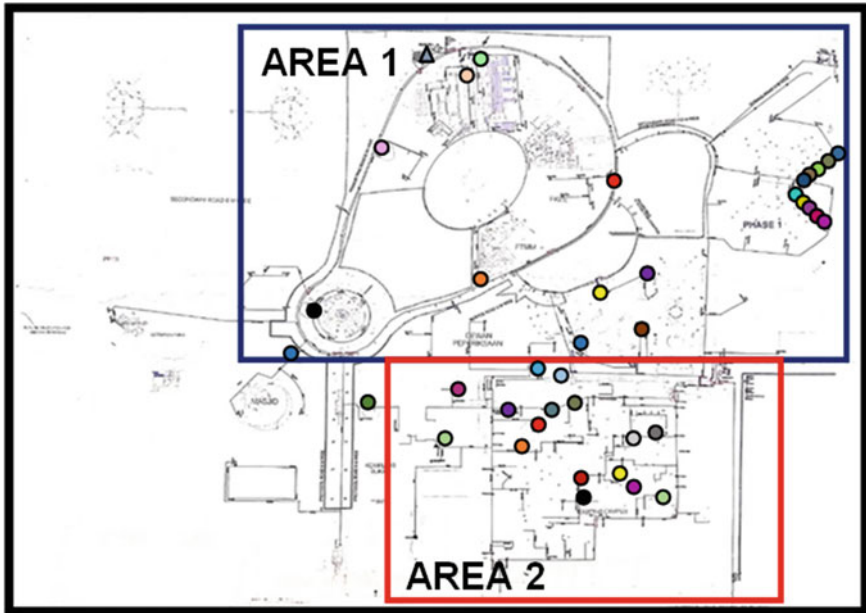


Fig. 37.1 Map of water reticulations system at Universiti Tun Hussein Onn Malaysia

37.2.1 *Finding Water Reticulation Map*

The water services map shows the known location of all water, storm-water and wastewater reticulation networks managed by Universiti Tun Hussein Onn Malaysia management. The map is intended specifically to identify main pipes and dead end of the pipe. Figure 37.1 shows the map of water reticulation system for Area 1 and 2 at Universiti Tun Hussein Onn Malaysia.

37.2.2 *Determining Flows Based on Metering*

Water meters measure the amount of water used by housing units and each buildings unit delivered with water by the municipal water supply scheme. They are also used to determine the flow across a specific section of the network. Flows can be easily detected by using water meters. The volumes are detected by reading the meter capsule from each meter. Bulk and volumetric water meters measure the flow rate of the water and record the total volumes that can subsequently be utilized for invoicing. The water supply network is divided into two areas which are known as Area 1 and 2. Area 1 is the new metering system where all the meters are located at the faculty, mosque, library, and G3 areas. Area 2 m are known as old zone and

old metering system where all the meters are located near blocks A, B, C, D, and E (Figs. 37.2 and 37.3).



Fig. 37.2 Bulk meter capsule reading



Fig. 37.3 Volumetric meter capsule reading

37.2.3 Water Balance Method of Calculation

Water balance calculation is the water balance and the water loss indices proposed by IWA are an important method for measuring and forecasting water losses. The following Eq. (37.1) is used to measure the percentage of losses.

$$WL(\%) = \left(\frac{SIV - BAC - RL}{SIV} \right) \times 100\% \quad (37.1)$$

where;

- WL Water loss (%)
- SIV Water supplied to the network (m³/week)
- BAC Water distributed to service pipe (m³/week)
- RL Real losses (10%)—fixed value

37.2.4 Analysis of the Water Losses in AWWA Dataset

A water audit is an analysis of all the water in the water system resulting in a quantified view of the quality and function of the water system. It is the first step in formulating an economically sustainable strategy to fix water losses. The preliminary audit of water starts with the following data and quick calculation. The following Eq. (37.2) is used to measure the current annual real losses.

- a. Determine the volume of water to be added to the system by reading bulk and volumetric meter.
- b. Determine the allowed consumption (billed + unbilled).

$$CARL = SIV - UAC \quad (37.2)$$

Where;

SIV is the water supplied to the network (m³/year),

UAC is the water used for own needs (m³/year).

Data Load in AWWA Software

AWWA developed the Water Audit Program, which was used in the current study. The program has been installed in an impressive Microsoft Office with a wide range of worksheets. When the data entry is complete, the application estimates spending, income and NRW and a comprehensive water balance [10]. The IWA and AWWA collaborated on a water audit system, which AWWA supports. The IWA / AWWA Water Audit System provide water utilities with best practice management tools and

recommendations for properly managing resources. The guidance to use the software are as follows:

- a All water system identifying and contact information must be entered on the “Instructions” page.
- b All water volumes must be entered in Million Gallons (MG).
- c Carefully read every condition, pick the one that fits the conditions most closely on the device, and enter the number of that condition in the box.
- d Validity scores are heavily weighted on the conditions in the “Water Supplied” section.
- e It is extremely important to check the master meter for accuracy and calibrate annually.

Tabulation of Data

Tabulating is a method of processing data or information by arranging it in a table. Values were entered in accordance with the AWWA software’s requirements. The following are the requirements for entering values into the software:

- a. Volume from own sources
 - Amount of treated water input into the system from own processing facilities
- b. Billed metered consumption
 - Both the meter utilization that is paid for and extends to all types of customers, such as residential, commercial, industrial, or institutional and does NOT include water sold to nearby services.
- c. Length of mains
 - Duration of all pipelines (except utility connections) in the system beginning from the input point of the system (for example at the outlet of the treatment plant)
- d. Number of active AND inactive service connections
 - Bulk water sold and conveyed out of the water distribution system. Typically, this is water sold to a neighboring water utility.
- e. Average operating pressure
 - When compiling a preliminary water audit, the average pressure can be approximated.
- f. Total annual cost of operating the water system
 - Such expenses include those for administration, repairs and any annual costs accrued for the long-term maintenance of the scheme, such as the redemption of capital bonds for infrastructure expansion or development.

Verification of Data

The water audit gives an initial calculation of the total number of real losses, but more leakage data and research are needed to allow for the cost-effective reduction of real losses. All the data is compared to each other in order to determine the most appropriate remedy for leakage, the utility can complete a section analysis of actual damage after the discovery and prevention of the water audit has been done.

Percentage of Losses

In order to calculate non-revenue water as a percentage of operating system spending, the audit software first calculates the volume of evident losses using the customer retail unit figure. The audit software then calculates the amount of real damages and unbilled authorized use using either the variable production cost (default) or the customer retail unit cost, based on the auditor's preference of costs.

37.3 Result and Discussions

The water balance in the study area was obtained from week 1 to week 7. Percentage of losses was calculated on the basis of the equation explained in the methodology. Values are estimated on the basis of data collected per day for 7 weeks. Real losses are measured on the basis of the world water assertion and the overall cost of losses is therefore determined on the basis of water consumption per m³.

37.3.1 Analysis of Water Balance Method

Tables 37.1 and 37.2 show the result of the overall amount of water measured per day for 7 weeks using water balance method. The total percentage of Area 1 losses

Table 37.1 Result for water loss in 7 weeks (Area 1)

Week	Water loss (%)	Water loss (RM)	Volume of water loss (m ³)
1	2.34	51,546.00	17,182
2	2.62	58,116.00	19,372
3	3.64	82,086.00	27,362
4	3.56	80,220.00	26,740
5	2.88	64,929.00	21,643
6	2.89	65,451.00	27,362
7	2.92	66,048.00	26,740
	Σ Total = 20.85	Σ Total = RM 468,396	156,132 m ³

Table 37.2 Result for water loss in 7 weeks (Area 2)

Week	Water loss (%)	Water loss (RM)	Volume of water loss (m ³)
1	31.21	472,929.00	157,643
2	31.20	472,740.00	157,580
3	31.18	472,479.00	157,493
4	31.14	471,945.00	157,315
5	31.13	471,933.00	157,311
6	31.13	472,017.00	157,339
7	31.12	471,975.00	157,325
	\sum Total = 218.11	\sum Total = RM 3,306,018	1,102,006 m ³

is 20.85% and according to the World Bank Group Area 1 water loss is below 25% for 7 weeks. This notes that there is no shortage of water. The overall percentage of Area 2 loss in Table 37.2 is 218.11% for 7 weeks, which is more than 25%. This has arisen because the infrastructure of water supply is inadequate without water meters. Some of the buildings did not provide an efficient water meter so that the total water distributed to the network is unidentifiable. Water meters have had too many flaws which are unidentifiable water meters, damaged meters, restricted area and other defective features because the meter provided in Area 2 was very old so the meter was inaccurate and it affects the result. So, this proves that there is no water loss due to any leakage but because of the damaged meters the volumes of water at certain places are unidentifiable.

Figures 37.4 and 37.5 show the comparison of the percentage of water losses collected over 7 weeks. The percentage of losses in Area 1 increased from 2.34 to 3.64% from week 1 to week 3. Then the percentage of losses dropped from 3.65 to 2.88% from week 3 to week 5. From there, the proportion had increased again from 2.88 to 2.92%, with a slight difference from week 5 to week 7. The percentage of losses in Area 2 fell slightly from 31.21 to 31.13% from week 1 to week 5 and then remained level for the following week. Then, the percentage of loss dropped again from week 6 to week 7 from 31.13 to 31.12%. There might not be a huge difference between the drop and the rise in the percentage because there is no any state of leakage has been found. Thus, the efficiency of the meter is good enough for the proper supply of water at all stations.

Analysis of AWWA Software

The analysis of AWWA software showed total loss of water. Both values have been obtained from the management of UTHM for water use per year. This methodology is utilised to show that there is another way to effectively discover losses by percentage, because the computer can classify annual water losses without using any mathematical methods to prove water losses. However, in order for this strategy to continue with the result, some understanding of the software and values to identify the length of main and water pressure would be required. The total volume of water will be

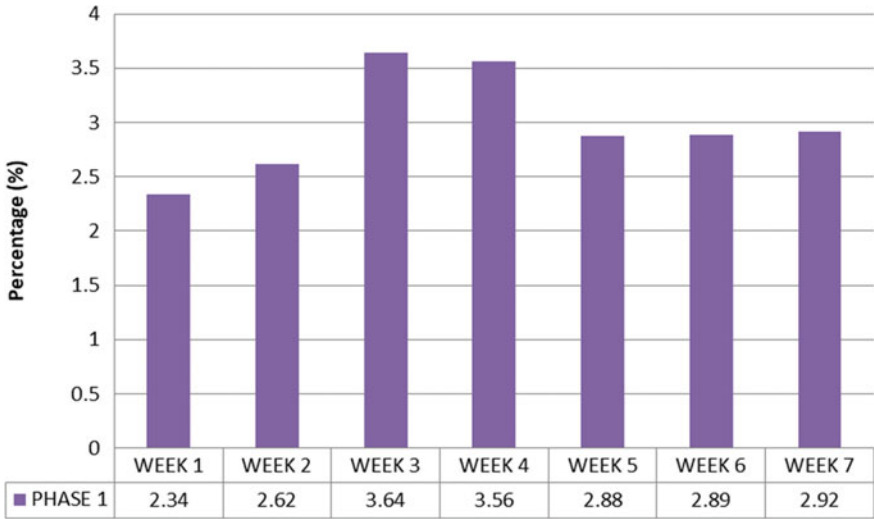


Fig. 37.4 Percentage of water loss for Area 1

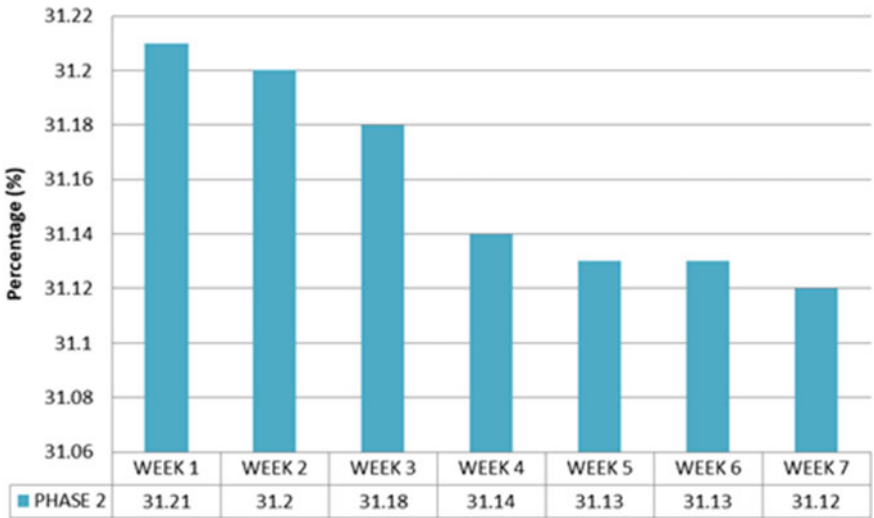


Fig. 37.5 Percentage of water loss for Area 2

converted to a million gallons (US). The data collected from Table 37.3 become input for the AWWA software to find the percentage per annual.

The first value most needed for the programme is the amount from its own sources, which is the total water distributed in the water supply system in Table 37.4. The total water supply had been recorded by the programme itself.

Table 37.3 Total water consumed in year 2020

Month	Amount (RM)	Volume of water (m ³)
January	263,666.82	78,910.00
February	268,291.30	81,311.00
March	241,323.90	72,878.00
April	218,407.32	66,195.00
May	208,554.70	63,209.00
June	179,749.00	54,480.00
July	202,356.11	63,531.00
August	168,851.22	51,178.00
September	193,955.50	58,785.00
October	192,340.62	58,296.00
November	169,550.23	20,167.00
December	171,100.50	24,316.00
Σ Total	RM 2,478,147.22	693,256 m ³

Table 37.4 Water supplied from own sources

Unit: m ³ convert to (MG/year)	
Description	Answer
Volume from own sources	734,396.00 m ³ convert to (MG/year) = 161,522,539 (MG/year)
Σ Water supplied	= 161,522,539 (MG/year)

The Authorized Consumption in Table 37.5 is identified by adding the value of the billed metered, defined as the total metered volume of water, from the data recorded each month. Unbilled unmetered was automatically incorporated into the software.

The length of the mains pipe as shown in Table 37.6 was determined from water reticulation map by calculating the length of the pipe. Then, lengths of the pipe multiply with the drawing scale and converting the unit to miles. The number of active and inactive service links is also determined by measuring the total number of stations on the map.

Table 37.5 Authorized consumption

Unit: m ³ convert to (MG/year)	
Description	Answer
Billed metered	693,256 m ³ convert to (MG/year) = 152,602,786 (MG/year)
Authorized consumption	= 154,621,817.739 (MG/year)

Table 37.6 System data

Unit: m ³ convert to (MG/year)	
Description	Answer
Length of Mains	= 310 cm × 4000 (scale of map) 1240000 cm convert to miles = 7.7 miles
Number of Active and Inactive service connections	= 39 connections
Average Length of customer service line(Pipe length between curb stop and customer meter)	= 12ft *Referred Water Reticulation Map
Average Operating Pressure	= 50psi *Determined from the main pressure water meter

*Percentage of loss from AWWA software analysis = 5.5% out of 25% (water loss limit)

37.4 Conclusion

According to the data, the average value of Area 1 water losses on campus is less than 25%, or 20.85% for seven weeks, and the value is lower by international standards. The findings also explained the key issues that occurred in Area 2, where the water loss value was greater than 25%, or 218.11% in just 7 weeks, but the causes were not due to water loss, but rather to unidentified water meters, damaged meters, a restricted area, and other flaws in the water balance method. Aside from that, the results showed that both the water balance technique and AWWA software are the best ways to find the percentage of water loss that is 5.5% less than 25% according to international standards, and that AWWA software helps to find the losses in such a way that no additional calculations are required, but that the software requires annual data for data tabulation. The proportion of water loss was determined using the water balancing method, and the effectiveness of the 5.5% on AWWA software for the purpose of water loss research was also examined. The proportion of water loss should not exceed 25%, according to international standards. As a result, there is no water shortage on campus. The loss of water is not the reason of high.

Acknowledgements The authors would also like to thank the Faculty of Civil and Built Environment and Pejabat Pembangunan dan Penyelenggaraan, Universiti Tun Hussein Onn Malaysia for its support.

References

1. T. AL-Washali, S. Sharma, M. Kennedy, F. AL-Nozaily, M. Haidera, Monitoring the non-revenue water performance. *Intermittent Supply Water* **11**(6), 1220, 1–14 (2019)
2. M. Farley, S. Trow, *Losses in water distribution networks; a practitioner's guide to assessment, monitoring, and control*, 2nd edn. (International Water Association Publishing, UK, 2013)
3. O. Hunaidi, A. Wang, A new system for locating leaks in urban water distribution pipes. *Manage. Environment. Quality: Int. J.* **17**, 450–466 (2006)
4. F.J. Arregui, E. Cabrera, R. Cobacho, J. Garcia-Serra, Reducing apparent losses caused by meters inaccuracies. *Water Practice Technol.* **1**(4), 1–8 (2006)
5. V. Aravinthan, G. Ho, J. Gray, R. Butler, D. Connell, What are smart meters?, 1–4 (2012)
6. R.M. Willis, R.A. Stewart, K. Panuwatwanich, S. Jones, A. Kyriakides, Alarming visual display monitors affecting shower end use water and energy conservation in Australian residential households. *Resour. Conserv. Recycl.* **54**(12), 1117–1127 (2010)
7. C. Drubin, Smart meter rollouts from water utilities gain momentum (2016)
8. G. Oren, N.Y. Stroh, Mathematical model for detection of leakage in domestic water supply systems by reading consumption from an analogue water meter. *Int. J. Environment. Sci. Develop.* **4**(4), 386–389 (2013)
9. T.C. Britton, R.A. Stewart, K.R. O'Halloran, Smart metering: enabler for rapid and effective post meter leakage identification and water loss management. *J. Clean. Prod.* **54**, 66–176 (2013)
10. M. Mastaller, P. Klingel, Adapting the IWA water balance to intermittent water supply and flat-rate tariffs without customer metering. *J. Water, Sanitation Hygiene Develop.* **7**(3), 396–406 (2017)

Chapter 38

Short-Term Time-Series Forecasting by Using Exponential Smoothing Techniques at Sungai Pahang, Malaysia



Ahmad Murshid Hamidon and Sabariah Musa

Abstract The increasing of river water level usually happens during raining season and can lead to devastating flash floods. Therefore, forecasting river water level series using the exponential smoothing method was applied to predict accurate river water level series. Three exponential smoothing techniques have been investigated to study their ability in handling extreme river water level time series data, which are Single Exponential Smoothing Technique, Double exponential smoothing technique and Holt's Method. The techniques were performed on river water level data from three rivers in Pahang, Malaysia. This method of forecasting is evaluated to determine the ability in the forecasting river water level for short-term forecast with seasonal and non-seasonal data. Based on the error generated from the analysis, Simple exponential smoothing technique from case study 1 was found to be the best model smoothing technique as it produced the lowest MAPE error which is 0.09% as it suitable for short-term forecasting in 6 months ahead. The selection of seasonal data in cases studies 2 and 3 while non-seasonal data in case study 1 also showed different situations in the forecasting results. More accurate forecasts can be created by identifying the appropriate smoothing approach for extreme data. A good prediction will help the government and the public to reduce the impact from flood catastrophes and to provide the public information about forthcoming events as an early warning system.

A. M. Hamidon · S. Musa (✉)

Department of Civil Engineering, Faculty of Civil Engineering and Built Environment, Universiti Tun Hussein Onn Malaysia, 86400 Batu Pahat, Malaysia
e-mail: sabariah@uthm.edu.my

S. Musa

Faculty of Civil Engineering and Built Environment, Eco-Hytech Research Centre, Universiti Tun Hussein Onn Malaysia, 86400 Batu Pahat, Malaysia

38.1 Introduction

38.1.1 Background

There is about 90% of water resources in Malaysia are used for industrial, domestic use and daily life purposes [1, 2]. The main source of water in Malaysia is mainly river water that comes from rainfall. Though the rainfall is considered abundant, flood problems often occurred in Malaysia due to the abundant rainfall in rivers in Malaysia. Natural disasters that often occur in Malaysia are floods and it is considered a continued hazard for humanity. One of the major causes of floods in Malaysia is the high and continuous distribution of rainwater from heavy rainfall for days. If this condition continues it can cause the river water level to rise drastically, this incident can cause devastating flash floods [3].

Sungai Pahang is one of the areas that received the highest total rainfall throughout the year. Therefore, the study of the prediction of river water level time series is very important in order to avoid flood events. Extreme event time series are difficult to study and even harder to be used for prediction because of their rare characteristics [4]. The exponential smoothing techniques applied on monthly river water level and identifies that can forecast and analyse Sungai Pahang monthly river water level time-series data for short-term forecast. Sungai Pahang was selected in this study because Sungai Pahang was the largest water source in peninsular Malaysia and flood was very often occurred in Sungai Pahang. For that reason, 3 stations in Sungai Pahang river were examined in this study. The basin of Sungai Pahang has an annually precipitation of approximately 2170 mm, much of which takes place between mid-October and mid-January during the North-East Monsoon [5]. The advantages of adopting exponential smoothing techniques include their ease of learning and application, as well as their ability to create accurate forecasts. A forecast for the next period is generated using the exponential smoothing approach. Forecasts for longer time periods can then be created using the trend projection technique. Finally, it lends greater weight to recent observations.

38.1.2 Overview on Exponential Smoothing Techniques

Three smoothing techniques which are Single Exponential Smoothing Technique (SEST), Double Exponential Smoothing Technique (DEST) and Holt winter are discussed.

Single Exponential Smoothing Technique

This model is ideal for series in which the pattern or seasonality is not present. The only parameter for smoothing is level. Single exponential smoothing is equivalent to an ARIMA Model with zero auto regression orders, one differentiation order, one average moving order, and no constant order. The Single exponential smoothing is

often referred to as simple exponential smoothing [1]. This technique is a common statistical method used for the prediction of time series and can also be used in pre-processing data [6]. Single exponential smoothing offers a simpler, less expensive, and easier-to-understand value [7]. In Single exponential smoothing there is only one smoothing parameter. This smoothing scheme starts by converting S_1 , to y_1 with:

In the third time period,

$$S_3 = \alpha y_{t-1} + (1 - \alpha) S_2 \tag{38.1}$$

where;

S = smoothing observation data, y = actual observation data, $t = 1, 2, \dots, n$.

It's not S_1 series is smoothed by the second observation version. For the time period t , the smoothed value of S , is found in the form (38.2):

$$S_t = \alpha y_{t-1} + (1 - \alpha) S_{t-1} \quad 0 < \alpha \leq 1 \quad t \geq 3 \tag{38.2}$$

This is the basic formula for exponential smoothing and the fixed formula or parameter α is known as the smoothing constant. This alternative can also be proved by substituting y_{t-1} into the base formula with y_t , is a recent observation.

The forecasting formula is the basic equation that is (38.3):

$$S_t = \alpha y_t + (1 - \alpha) S_t \quad 0 < \alpha \leq 1 \quad t > 0 \tag{38.3}$$

It can be written as see (38.4):

$$S_{t+1} = S_t + \alpha(\epsilon'_t) \tag{38.4}$$

where; e is the prediction error (actual prediction) in time t . In other words, the most recent prediction is a value adjusted for errors made in past predictions.

Double Exponential Smoothing

Additionally, referred to as Brown's method [8]. It is advantageous for time series with a linear trend. The primary advantage of double exponential smoothing is its capacity to make multiple-ahead forecasts, yet the value of is notoriously difficult to quantify α [9].

This scheme can be proved by introducing a second formula with a fixed value, y' , which should be chosen to be combined with α . Here are two Eqs. (38.5) and (38.6) presented with Double Exponential Smoothing.

$$S_t = \alpha y_t + (1 - \alpha) (S_{t-1} + b_{t-1}) \quad 0 \leq \alpha \leq 1 \tag{38.5}$$

$$b = y'(S - S) + (1 - y) b \quad 0 \leq y' \leq 1 \tag{38.6}$$

where is the latest from the series is used to calculate the smoothed value to be replaced into multiple exponential smoothing.

In the case of Simple Exponential Smoothing, there are various schemes to provide the initial values of S_t and b_t for Multiple Smoothing. S_t is the basic preparation for y_t . Here are three types of suggestions for b_t see Eqs. (38.7), (38.8) and (38.9):

$$b_t = y_2 - y_1 \quad (38.7)$$

$$b = [(y_2 - y_1) + (y_3 - y_2) + (y_4 - y_3)]/3 \quad (38.8)$$

$$b_t = (y_n - y_1)/(n - 1) \quad (38.9)$$

There is also a forecast formula for a one -period ahead given below:

$$S_{t+1} = S_t + b_t \quad (38.10)$$

For the forecast of n-periods-ahead the following formula is given:

$$S_{t+n} = S_t + nb_t \quad (38.11)$$

Holt's Winter Model

This method is nearly the same as the simple exponential smoothing method, but has the advantage of reducing the update period of the trend component. The value of the data is a smoothed estimate of the value of the data at the end of each period, and the growth in the data is a smoothed estimate of average growth in the data at the end of each period [10]. The model is based on three variables which consist of see eqs. (38.12), (38.13) and (38.14):

The level estimate

$$L_t = \alpha y_t + (1 - \alpha)L_{t-1} = (L_{t-1} + T_{t-1}) \quad (38.12)$$

The trend estimate

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1} \quad (38.13)$$

Forecast m period into the future

$$Y_{t+m} = pTt + Lt \quad (38.14)$$

where; L_t = New smoothed value, α = Smoothing constant for the level, Y_t = Real value of the series in period t (actual value), β = Smoothing constant for trend

estimation, T_t = Trend Estimate, p = Period to be predicted, Y_{t+p} = Forecast for period p (estimated values).

The value of α , β and γ is fixed to be estimated in the MAPE which has minimal error. To start the Holt Winter method requires at least sufficient seasonal data for the beginning of estimates from seasonal indices I . Sufficient seasonal data indicates its period, L to estimate the trend factor from one period to the next. To complete it, it is necessary to use two seasonal data i.e. period, $2L$.

38.2 Methods

The methods section, otherwise known as methodology, describes all the necessary information that is required to obtain the results of the study.

38.2.1 Case Study

The data were collected and gathered from the Department of Irrigation and Drainage of Malaysia. The data used for experiment and testing were the historical data of river water level from January 2010 to February 2021. These data were collected in Pahang from three rivers of Sungai Pahang which were Sungai Jelai in Jeram Bungor Kuala Lipis as case study 1, Sungai Tembeling in Kampung Merting as case study 2 and Sungai Pahang in Temerloh as case study 3. The total number of time-series monthly river water level data was obtained from Department of Irrigation and Drainage Malaysia (DID) selected for experiment and testing is 122 series data.

38.2.2 Result Error

The results were simulated using the Exponential Smoothing Technique. First, the data is used to determine whether trends, seasonality, or both were included in the model. The performance evaluation of techniques was evaluated based on error measurement obtained by using performance metrics.

These measurements are based on the forecast errors, or how different the actual forecast is compared to the forecast. To test the model's ability to make accurate predictions, we utilise the Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE). The equations are as follows:

Percentage Error (PE)

$$PE = \frac{y_1 - \hat{y}_1}{y_1} \times 100 \quad (38.15)$$

Mean Absolute Percentage Error identifies significant relationships between forecast data and actual monthly water level.

$$MAPE = \frac{1}{n} \sum_{t=1}^n \{PE\} \tag{38.16}$$

Root Mean Squared Error is used to measures the differences between fitted value and actual.

$$RMSE = \sqrt{\frac{\sum_{t=1}^n e_t^2}{n}} \tag{38.17}$$

MAPE identifies significant relationships between forecast data and actual monthly water level. Models with a MAPE of around 30% produce reasonable predictions while MAPEs between 5 and 10% produce very accurate predictions. Analyzes that produce MAPEs of around 5 to 10% can be considered as accurate predictions.

To determining the value of the smoothing parameter, α is based on trial and error. In order to find the best value, the value of, α were tested which begin with 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, and 0.9. The α that gives the smallest error will defined as the best value.

38.3 Result and Discussions

The results and discussion section presents data and analysis of the study.

38.3.1 Results

The tentative models developed are used to estimate its parameters as scheduled in Tables 38.1, 38.2 and 38.3 are produced giving SSE (Sum of Squared Errors), MAPE and RMSE values respectively.

Table 38.1 Result comparison using Sungai Jelai dataset

Model	SSE	Parameter	MAPE (fit)	MAPE (forecast)	RMSE
SEST	57.00	$\alpha = 0.6$	0.441	0.280	0.490
DEST	76.90	$\alpha = 0.2$	0.517	0.227	0.522
Holt winter	41.22	$\alpha = 0.6$ $Y = 0.0$	0.557	0.323	0.441

Table 38.2 Result comparison using Sungai Tembeling dataset

Model	SSE	Parameter	MAPE (fit)	MAPE (forecast)	RMSE
SEST	44.53	$\alpha = 0.93$	3.273	2.913	3.115
DEST	142.348	$\alpha = 0.8$	4.635	0.359	3.428
Holt winter	19.48	$\alpha = 0.5$ $Y = 0.0$	2.713	1.304	2.135

Table 38.3 Result comparison using Sungai Temerloh dataset

Model	SSE	Parameter	MAPE (fit)	MAPE (forecast)	RMSE
SEST	41.259	$\alpha = 0.48$	3.535	1.643	1.594
DEST	44.732	$\alpha = 0.1$	3.885	0.838	1.637
Holt winter	38.240	$\alpha = 0.5$ $Y = 0.0$	3.116	1.511	1.207

Referring to Table 38.1, the MAPE on the forecast data and fit values data generated by the SEST model is lower than the other. If seen in Table 38.1, also the RMSE for adaptation data shows the second lowest is SEST model. SEST model can be selected as the best model, since the SEST model has the second lowest SSE and RMSE values. Based on Table 38.2, the MAPE values on the forecast data and fit values were generated by the Holt winter model is lower than the other. However, all MAPE values of those models are less than 10%. But between this three model Holt winter has the lowest value which is 0.3%. Even so, with the RMSE also in Table 38.2 shows Holt winter model gives the lowest. As such, the Holt winter model was chosen as the best model in this case. Based on Table 38.3, the MAPE values on the forecast data and fit values were generated by the Holt winter model is lower than the other. However, all MAPE values of those models are less than 10%. But between this three model Holt winter has the lowest value which is 0.42%. Even so, with the RMSE also in Table 38.3 shows Holt winter model gives the lowest which is 1.207. As such, the Holt winter model was chosen as the best model in this case.

The time series plots in Figs. 38.1, 38.2 and 38.3 compare actual data to smoothed data using DEST, SEST, and Holt’s method for Sungai Jelai in Jeram Bungor, Sungai Tembeling in Kampung Merting, and Sungai Pahang in Temerloh, respectively. The x-axis represents the date, while the y-axis represents the monthly water level data in meters. According to Fig. 38.1, the values data obtained using SEST were found to be largely consistent with the actual data. This result demonstrates that SEST outperformed DEST and Holt’s technique. According to Fig. 38.2, the data derived from Holt winter was determined to be largely consistent with the actual data. This result indicates that the Holt winter model outperformed SEST and DEST. This demonstrates the suitability of the Holt winter models for forecasting. According to Fig. 38.3, the fit values for data using Holt winter were determined to be largely consistent with the actual data. This result indicates that the Holt winter model

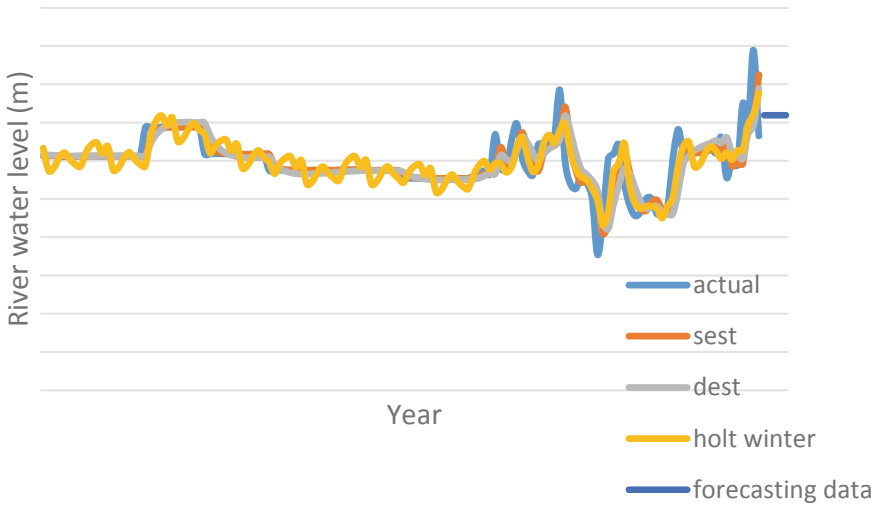


Fig. 38.1 Time-series plot for Sungai Jelai

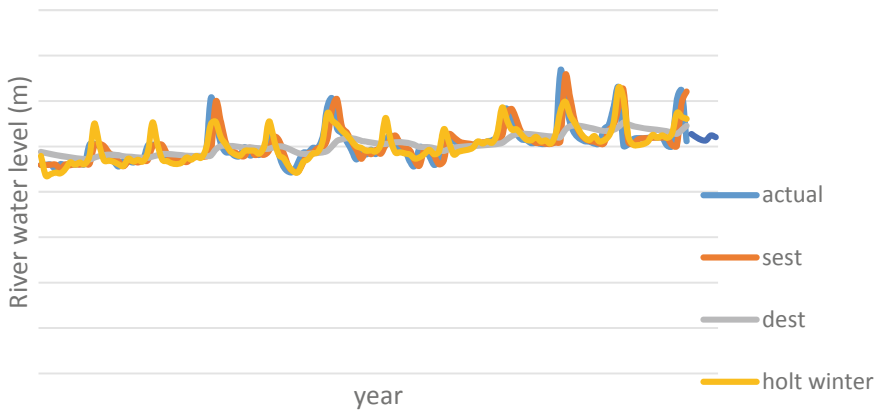


Fig. 38.2 Time-series plot for Sungai Tembeling

outperformed SEST and DEST. This demonstrates the suitability of the Holt winter models for forecasting.

38.3.2 Discussions

Based on Table 38.4, if we look at study case 1, the three models find that the MAPE produced is lower than studies case 2 and 3. It is likely that the monthly river water

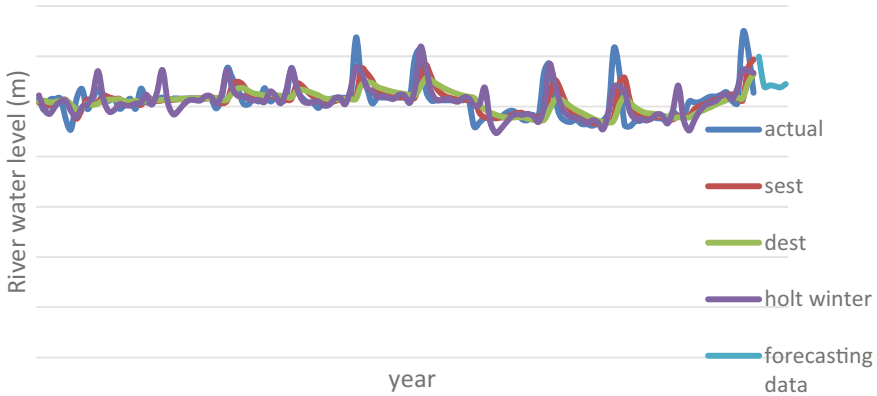


Fig. 38.3 Time-series plot for Sungai Temerloh

Table 38.4 The value comparison result of exponential smoothing technique

Case study	Model	SSE	MAPE (%)	RMSE	Parameter
Case study 1: Sungai Jelai in JeramBungor	SEST	57.00	0.09	0.490	$\alpha = 0.6$
	DEST	76.90	0.46	0.522	$\alpha = 0.2$
	Holt winter	41.22	0.34	0.441	$\alpha = 0.6$
Case study 2: Sungai Tembeling in Kg Merting	SEST	44.53	0.43	3.115	$\alpha = 0.93$
	DEST	142.35	0.76	3.428	$\alpha = 0.8$
	Holt winter	19.48	0.3	2.135	$\alpha = 0.5$
Case study 3: Sungai Pahang in Temerloh	SEST	41.30	0.44	1.594	$\alpha = 0.48$
	DEST	44.73	0.64	1.637	$\alpha = 0.1$
	Holt winter	38.24	0.42	3.116	$\alpha = 0.5$

level time series data of this study case is more uniform and does not contain seasonal variation components in the time-series. This case model of study 2 and 3 is also a simplified model with the lowest number of parameters.

The models that have been analysed need to be updated after having the latest data. To produce a good forecast, the model was reconstructed using all-time series data including new data every six months. This is because the Exponential Smoothing model forecasting methods are suitable for short-term forecasting with 6-month forecasting ahead.

Since this study analyses seasonal and non-seasonal series data, it gives the impression that the accuracy of the forecast results depends heavily on the appropriateness and number of time series data available. Similarly, the time data component greatly influences the shape of the river water level forecast direction.

38.4 Conclusion

This study has fulfilled the objectives of the study by developing and evaluating the model of Exponential Smoothing to predict the monthly water level of the river. All models developed are able to provide good accurate predictions. The simple exponential smoothing technique is the best model with the lowest MAPE value error of 0.09%. However, all models formed gave MAPE the lowest values because it has error below 5%. This study shows that the ability of the Exponential Smoothing methods to predict accurately is reasonable. The Exponential Smoothing model shows its ability to be more accurate for 6 months forecasting. The selection of seasonal data in cases study 2 and 3 while non-seasonal data in case study 1 also showed different situations in the forecasting results. The accuracy of the data is an important role in determining the accuracy of the forecast results. The amount of data must be sufficient and complete to form a good model. By finding the best smoothing technique for big data, more accurate prediction can be produced.

Acknowledgements The authors would like to thank the Faculty of Civil Engineering and Built Environment, Universiti Tun Hussein Onn Malaysia for its support. The author would also like to thank the Department of Irrigation and Drainage Malaysia for supplying hydrology and river water level data for this study.

References

1. K.Y. Chan, T.S. Dillon, J. Singh, E. Chang, Traffic flow forecasting neural networks based on exponential smoothing method, in *6th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, pp. 376–381 (2011)
2. S. Gelper, R. Fried, C. Croux, Robust forecasting with exponential and holt–winters smoothing. *J. Forecast.* **29**(3), 285–300 (2010)
3. P. Georgia, T. Hristos, Hydrological time series forecasting using simple combinations: big data testing and investigations on one-year ahead river flow predictability. Department of Water Resources and Environmental Engineering, School of Civil Engineering, National Technical University of Athens, HeroonPolytechniou 5, 157 80 Zographou, Greece (2020)
4. M. Ghil, P. Yiou, S. Hallegatte, B.D. Malamud, P. Naveau, A. Soloviev, I. Zaliapin, Extreme events: dynamics, statistics and prediction. *Nonlin. Processes Geophys.* **18**(3), 295–350 (2011)
5. A. Faisal, N.S. Azmin, A. Heshmatpoor, R. Hafiz, Identification of flood source areas in Pahang River Basin, Peninsular Malaysia. *Int. J.* **3**, 73–78 (2019)
6. R.J. Hyndman, G. Athanasopoulos, *Forecasting: Principles and Practice* (Monash University, Melbourne, 2018)
7. M.M.F. Hussain, R.A. Jamel, Statistical analysis & asthmatic patients in Sulaimaniyah governorate in the tuber-closes center. *Int. J. Res. Med. Health Sci.* **1**(2), 24–32 (2013)
8. M.A. Lazim, *Introductory business forecasting a practical approach*, 3rd edn. (UiTM Press, Shah Alam, 2012)
9. C.S. Jheison, L. Fernando, Analysis of energy consumption in Colombia using the Holt method. *Int. J. Energy Econ. Policy* **10**(6), 679–683 (2020)
10. P.S. Kalekar, Time series forecasting using holt-winters exponential smoothing, pp. 1–13 (2004)

Chapter 39

Internet of Things for Hybrid Energy System Data Monitoring



Saadatul Suhaili, Faridahanim Ahmad , Tsu Yian Lee ,
Nur Izieadiana Abidin, and Mariah Awang

Abstract Internet of things (IoT) refers to the sensors and electronic devices that are able to collect and share data over the internet. In remote areas where the electricity supply is limited, hybrid renewable energy systems are often developed to provide electricity. This paper aims to implement the IoT technology to monitor the energy efficiency of a Hybrid Energy System. The solar and mini-hydropower plants are the renewable energy sources of the Hybrid Energy System in this project. The Hybrid Energy System was developed to supply electricity to a guard house, while the IoT technology which consists of sensors and software was developed to record and analyze the electricity supply from the system in real time. By having the IoT-based system, the staff can properly monitor the pattern of electricity generation in real-time by using the software's user interface. This study focuses on the design of a user interface for the effective monitoring of energy generation data. Besides, a user acceptance test was implemented to make sure the software can fulfil user needs. In conclusion, with the aid of IoT technology, the energy generation of the Hybrid Energy System can be monitored properly in real-time, and the user can easily communicate with the system using the user interface of the software.

S. Suhaili · F. Ahmad (✉) · T. Y. Lee · N. I. Abidin
Department of Structure and Material, School of Civil Engineering, Faculty of Engineering,
Universiti Teknologi Malaysia, 81310 Bahru, Johor, Malaysia
e-mail: faridahanim@utm.my

T. Y. Lee
e-mail: leetsuyian@graduate.utm.my

N. I. Abidin
e-mail: izieadiana@utm.my

M. Awang
Department of Civil Engineering Technology, Faculty of Engineering Technology, Universiti Tun
Hussein Onn Malaysia (Pagoh Campus), Panchor Muar, 84600 Batu Pahat, Johor, Malaysia
e-mail: mariah@uthm.edu.my

39.1 Introduction

The IoT is a broad term referring to electronic devices and sensors capable of collecting information and communication, wired or wireless [1]. Basically, a complete IoT system consists of sensors, network connection, and software to achieve the designed purpose. Simply put, the targeted data collected by a specific sensor can be sent to and organized in a platform accessible by a human with internet access. IoT enables human to manage raw data more effectively. Hence, this technology has been successfully applied in different industry such as manufacturing industry, healthcare, and supply chain for different purpose [2].

IoT for Hybrid Energy System is a system that utilizes IoT technology to monitor the data of energy generation from the Hybrid Energy System remotely in real time [3, 4]. With the data of energy generation collected by IoT from the Hybrid Energy System, the total savings by day, month, and year can also be estimated. Hence, enabling the analysis of the efficiency of the Hybrid Energy System to be conducted.

In this study, the hybrid energy system is the system consisting of a mini hydro and six solar panels for power generation. The mini hydro was set up in the middle of the stream to transfer the wave energy to the hybrid power supply, while the solar panels located by the stream produce solar energy as additional power to the hybrid system [5]. The main purpose of developing the Hybrid Energy System in this study is to provide the electricity supply to the guard house nearby. Regardless of the design of the Hybrid Energy System, this study focuses on solving the issue of how the energy generation from the system can be recorded and analyzed in real time. This is because the data of energy generation is significant for inspection of the system to ensure it works smoothly daily. The use of manual way to record the data daily, monthly, and yearly on the site by the technician is considered inefficient and requires a real-time solution. In a conventional method, the technician collects the data daily and manually sends the data via email to the staff off-site. To overcome this problem, an Internet of Things (IoT) system was designed. In order to implement the IoT technology to monitor the energy efficiency of a Hybrid Energy System, few objectives were derived, which are to design an IoT-based software to monitor a Hybrid Energy System and to analyze the effectiveness of the Hybrid Energy Software.

39.2 Literature Review

39.2.1 *Hybrid Energy System*

Electricity is significant to support daily human activities such as powering electrical appliances and machineries. Meanwhile, carbon dioxide (CO₂) emission from the electricity generation industry due to the high dependency on fossil fuel as an energy source has been addressed [6]. In addition, due to the fluctuation of global economic

and fossil fuel prices, the use of renewable energy resources such as wind, solar, hydro, and biofuels for electricity generation is currently a hot topic that is being investigated extensively to reduce fossil fuel use. Besides being able to solve the environmental pollution problem, the use of renewable energy is also cost-efficient.

One issue of renewable energy resources is the unpredictable energy generation because it fully depends on the ever-changing environment condition [7]. In this case, Hybrid Energy System is an alternate solution. A hybrid energy system is the power generation system that uses two or more types of energy sources to generate power [8]. For example, mini-hydropower and solar Photovoltaic are combined to form a hybrid energy system to generate electricity. When designing the hybrid energy system, the meteorological and seasonal weather conditions must be taken into consideration to ensure the continuous supply of the electricity. Therefore, the energy generation operation of the hybrid system will not only depend on a single renewable resource.

In this paper, the Hybrid Energy System, which combines mini hydro system and solar panels, is installed to supply electricity to a guard house at remote area, nearby to a stream. Both kinetic energy from the mini hydro system and solar energy from the solar system are converted to electricity. Apart from optimizing the system design, monitoring the performance of the operation is required to ensure that energy is generated all the time as expected. To effectively monitor energy generation data, Internet of Things (IoT) adoption is one suitable solution.

39.2.2 Internet of Things in Hybrid Energy System

Internet of Things (IoT) refers to “the stringent connectedness between digital and physical world” [1]. IoT system works so that with internet connection, devices within the IoT infrastructure shares data among them, send data to a centralized server or cloud-based application, and store data in the database for further data processing and performing different designed tasks [1]. The user performs tasks in the cloud-based application with data collected from the IoT system.

IoT has been adopted in various fields, including Hybrid Energy systems. For example, the role of IoT in monitoring a campus energy generation by Hybrid Energy System has been discussed in [3]. An android based online web for energy generation from renewable energy system monitoring was also proposed by Jiju et al. [4]. Both papers highlighted that monitoring energy generation and cost savings from the Hybrid Energy System became more efficient when IoT was adopted. However, the discussions in the papers focus more on the system architecture, elements and communication protocols of the IoT system, without stressing the importance of data arrangement in the system user interface. In fact, the design of the user interface significantly determines the monitoring work effectiveness. Therefore, this study would like to emphasize the discussion on the user interface design of the software and the testing to ensure system effectiveness.

39.3 Methodologies

The methodologies of this study were designed to achieve the two objectives. Firstly, the user interface of the IoT software was developed. Then, the effectiveness of the software was investigated based on the user acceptance survey.

39.3.1 Software User Interface Design and Development

The diagram explaining how the Hybrid Energy System is monitored by the IoT system is as shown in Fig. 39.1. The Bluetooth device installed on the site of the Hybrid Energy System sends the energy generation data to the software database hosted in the cloud. Then, the staff or user monitors the data saved in the database of the data server by accessing the cloud, using a desktop with an internet connection. The user interface design of the system, which is the display of software on the desktop monitor is significant to ensure that it is user friendly and capable of fulfilling users' demands. Hence, the software's user interface was designed based on the user requirement investigated from the interview session in 3 months.

Since the software is designed for recording the power generation and calculating the cost savings from the Hybrid Energy System, there are some mathematical formulations applied in the software. Based on Ohms Law, the power generation can be calculated by simply multiply the current value and the voltage value as shown in the equation below:

$$P = I \times V \quad (39.1)$$

where,

- P Power (W)
- I Current (A)
- V Voltage (V)

For calculating the electricity cost savings, the rate (RM) of Block Tariff, provided by Tenaga National Berhad (TNB) as shown in Table 39.1 is referred. The rate for Block Tariff of less than 200 kWh and more than 200 kWh is different.

Table 39.2 shows how cost savings from the Hybrid Energy System can be calculated using the Tariff rate. Power input from the solar and mini hydro systems are multiplied with the Tariff rate to estimate the cost that can be saved from paying to TNB. The value of power input provided in Table 39.2 is only an assumption. In the developed software, the calculation is performed automatically using the programming algorithm, without needing manual calculation.

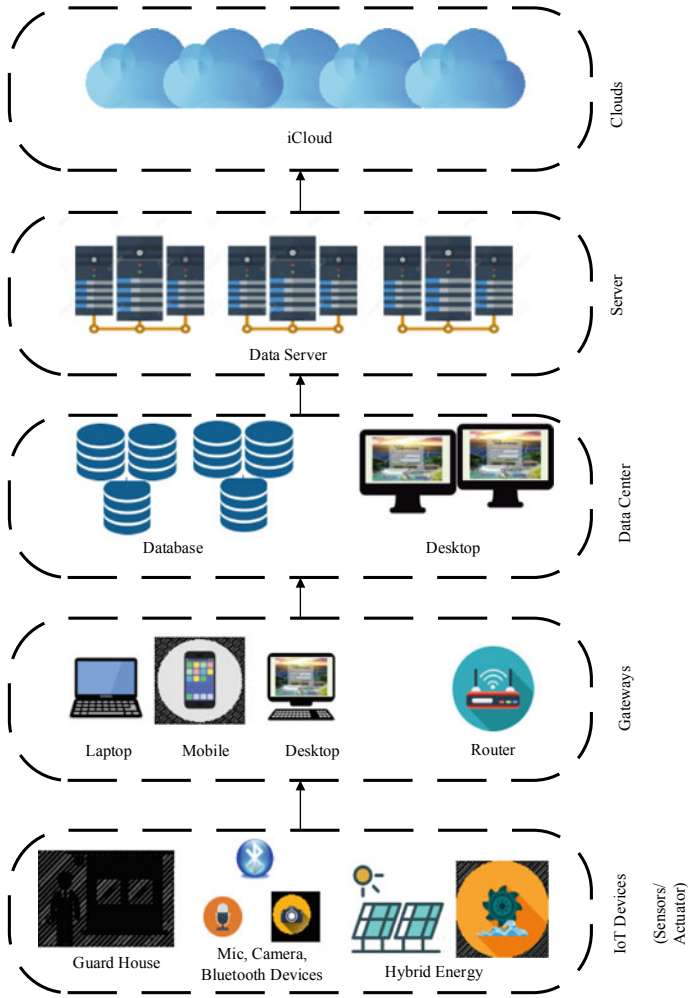


Fig. 39.1 Diagram of IoT in hybrid energy system

Table 39.1 Rate of block tariff

Block Tariff (kWh)	Prorate factor	Rate (RM)
200	31 Day/31 Day = 1.00000	0.435
> 200	31 Day/31 Day = 1.00000	0.509

Table 39.2 Estimation of cost savings with hybrid energy system implemented

Power generation source	Power input (kW)	Rate (RM)		Expected spend if no hybrid energy system installed (RM)	Cost savings from hybrid energy system (RM)
TNB	250	200	0.435	$0.435 \times 200 = 87.00$	-
		> 200	0.509	$0.509 \times 50 = 29.50$	
Solar panel	650	200	0.435	$0.435 \times 200 = 87.00$	316.05
		> 200	0.509	$0.509 \times 450 = 229.05$	
Mini Hydro	600	200	0.435	$0.435 \times 200 = 87.00$	290.60
		> 200	0.509	$0.509 \times 400 = 203.60$	
Total	1500	-		723.15	606.65

39.3.2 System Effectiveness Evaluation by User Acceptance Test

The software’s user interface serves as the interface for the user to communicate with the IoT system. Besides collecting real-time data of the Hybrid Energy System, the data is also displayed in an organized way that allows further analysis. Before the system is launched for use, the system needs to be evaluated by a user acceptance test. User acceptance test consists of Form-Based Test, User Security Matrix, User Acceptance Business Process Test Scripts and Report Test Script. The results of the user acceptance test are then summarized in the Defect Tracking Log.

39.4 Results and Discussion

39.4.1 IoT in Hybrid Energy System Software Design

The first objective is to design the software for IoT in Hybrid Energy. The system’s user interface or the module design is displayed in Figs. 39.2, 39.3, 39.4, 39.5, 39.6, 39.7 and 39.8. The log in page as shown in Fig. 39.2 is a page for the authorized person, usually staff, to log in to the system. Only the person allowed by the system admin can access to the system. All data in the system is controlled by the system admin. After inserting the username and password in the log-in page, the user will be

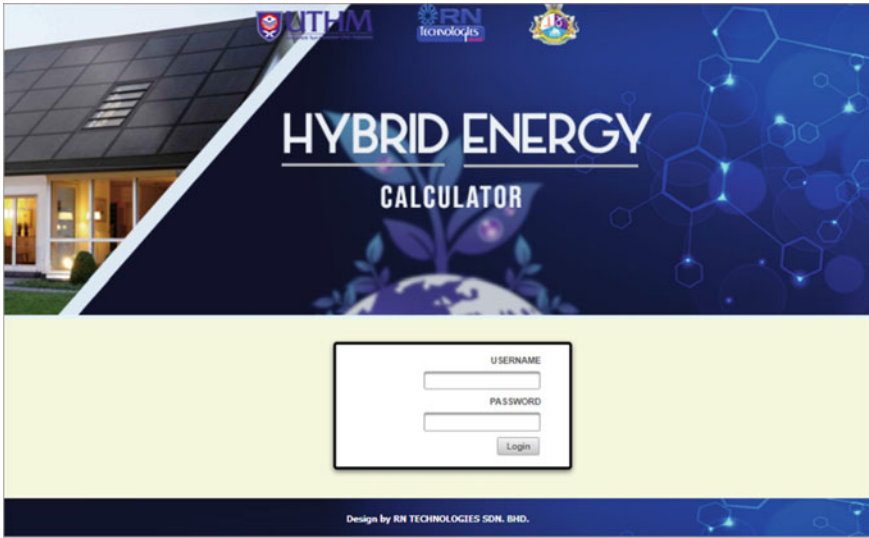


Fig. 39.2 Log in module

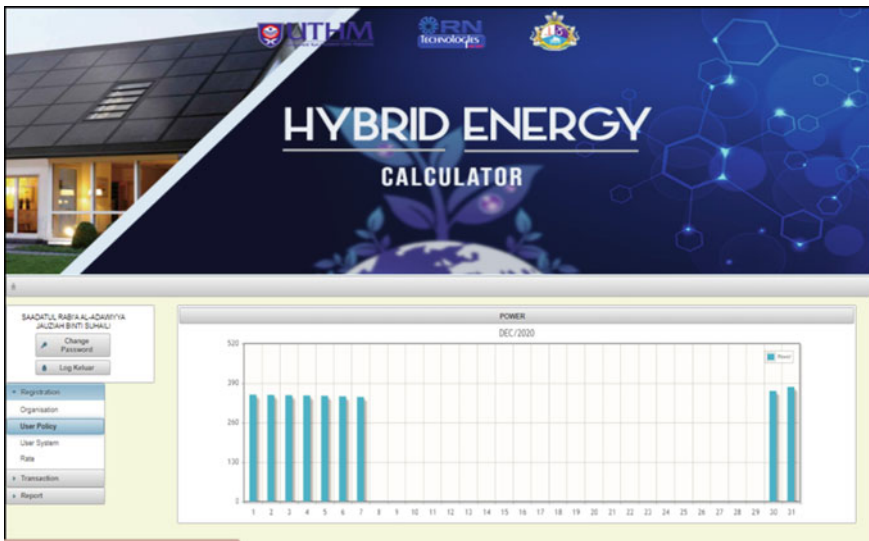


Fig. 39.3 Homepage module

led to the Homepage module (Fig. 39.3). The overall graph of energy generation from the Hybrid Energy System will be displayed at the Homepage module for evaluation.

Figure 39.4 shows the registration module, consisting of the sub-modules of Organization, User Policy, and User System (Fig. 39.5) and Rate module (Fig. 39.6). The

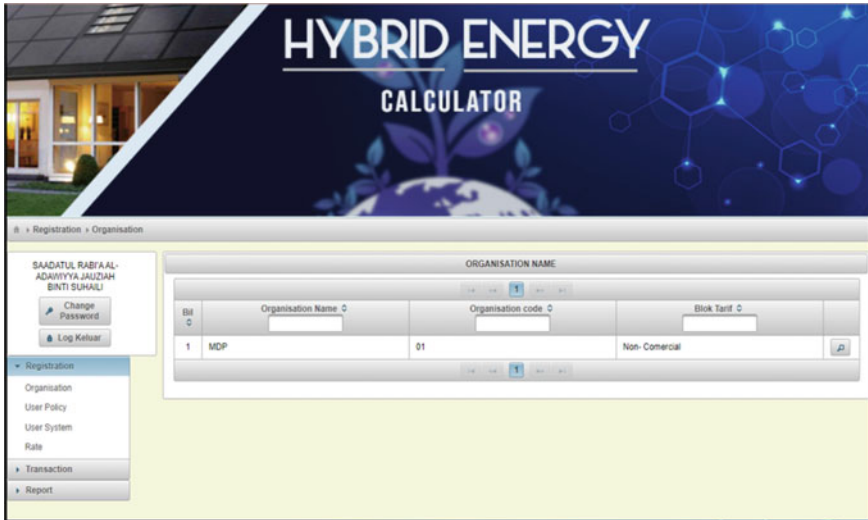


Fig. 39.4 Registration module

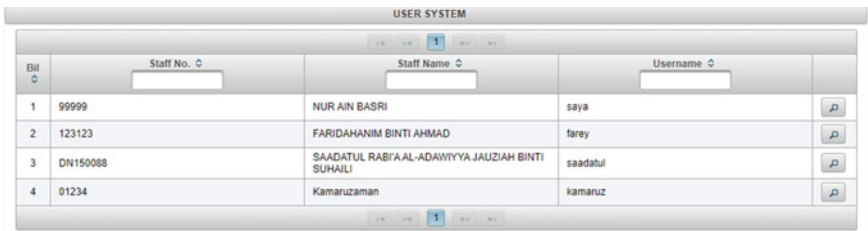


Fig. 39.5 User system module



Fig. 39.6 Rate module

organization can set the organization’s name, organization code, and the classification of block tariff using the Organization sub-module. The user system sub-module is recording the information of the authorized person, such as staff number, name, and username. Rate sub-module is used to set the Block Tariff name and Block Tariff description.

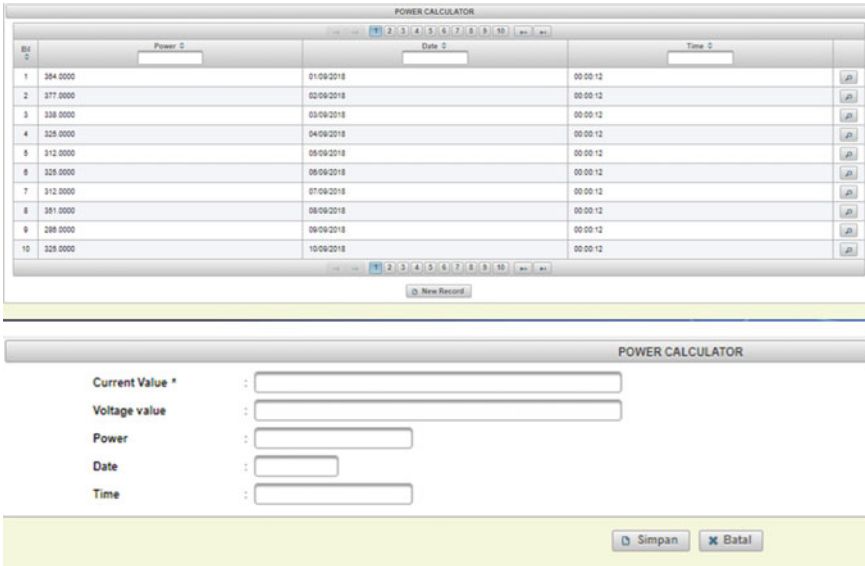


Fig. 39.7 Transaction and power calculator module

SOLAR POWER REPORT			
Bill	Date	Power	Amount (RM)
1	01/04/2019	377.0000	177.09
2	02/04/2019	390.0000	183.71
3	03/04/2019	377.0000	177.09
4	04/04/2019	364.0000	170.48
5	05/04/2019	390.0000	183.71
6	06/04/2019	377.0000	177.09
7	07/04/2019	351.0000	163.86
8	08/04/2019	364.0000	170.48
9	09/04/2019	377.0000	177.09
10	10/04/2019	390.0000	183.71
11	11/04/2019	390.0000	183.71
12	12/04/2019	364.0000	170.48

Fig. 39.8 Report module

The transaction and power calculator module as shown in Fig. 39.7 is used for the daily calculation of the system’s power generation. To calculate the power generation, the current, voltage, date and time are inserted into the power calculator. Then, the value of power will be automatically generated in the list and displayed in the graph at the Homepage of the system.

Figure 39.8 shows the daily report of power generation and total savings for electricity generation calculated, in table forms. The data generated can further be analyzed on a daily, monthly, and yearly basis to draw a meaningful conclusion.

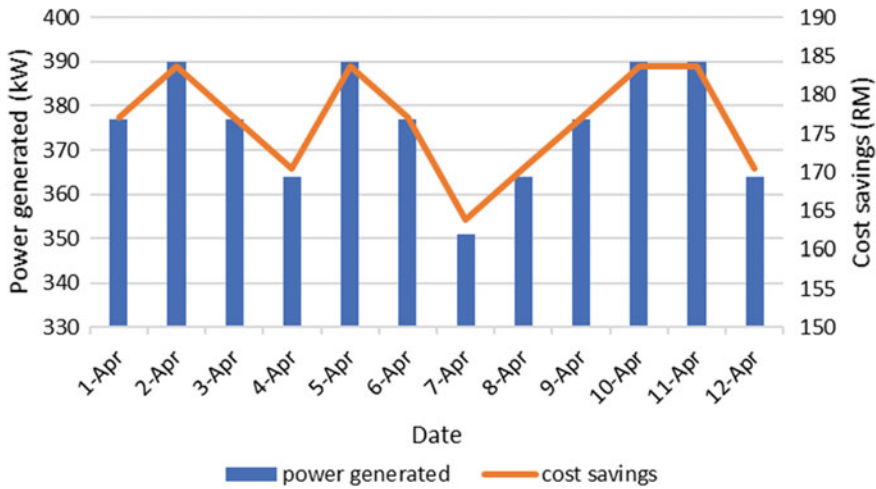


Fig. 39.9 The summary of power generation and cost savings report recorded in April

Figure 39.9 is also illustrated to demonstrate the graphical representation of the daily report in April 2019.

39.4.2 User Acceptance Test

The IoT system effectiveness is verified by conducting a user acceptance test before launching. All users performed the user acceptance test, with limitation to their works policy. For example, the system admin tested whether all information is accessible, and the clerk tested whether the data can be keyed into the system. The user interface design was corrected until the client was satisfied with the system. Meanwhile, the method to value the acceptance of the client is by conducting a user acceptance test. A defect tracking log was used as the instrument to evaluate the user acceptance to the system as shown in Table 39.3.

Based on Table 39.3, the defects of the system mentioned by the user were highlighted with the problem severity and the frequency of repeating problem. “E” refers to every time; “S” refers to sometimes; “O” occasionally; “1x” refers to happened once. Thereby, providing hints on how the system could be improved. After the improvement regarding the defects highlighted has been made, design review, equipment labelling, and visual inspection of the system were also implemented at the final stage of system development.

Table 39.3 Defect tracking log

Tester ID: Saadatul Date reported: 21 May 2019		User level tester: system admin Date report to contractor: 25/5/2019	
No.	Description	Severity (1,2,3,4,5)	Repeated (E,S,O,1x)
1	The items case sensitivity. (i.e., Do fields allow lower case text if they should only accept the upper case?) <ul style="list-style-type: none"> Uppercase and lowercase should be allowed for username and password. For security purpose, the password should mix upper and lower case text and number 	2	E
2	User Policy Module <ul style="list-style-type: none"> There is nothing shown in this module 	1	E
3	Power Calculation Module <ul style="list-style-type: none"> The user can hardly edit the data again after key in the data because the edit button was not functioning 	2	E
4	Report Module <ul style="list-style-type: none"> The report title should be written as “Hybrid Energy Report”, but not “Solar Power Report” 	1	E
5	Graph Module <ul style="list-style-type: none"> The graph that appears at the homepage should only show data for the latest month, but not mix with data for the previous month, which may be confusing 	1	E

39.5 Conclusion and Recommendation

In conclusion, this paper presents software design for IoT in Hybrid Energy System and the IoT system effectiveness testing. The IoT system is the platform for data analysis of energy generated from the Hybrid Energy System. The design of the system user interface was described by displaying few system modules, such as Registration Module, User System Module, Rate Module. Transaction and Power Calculator Module, and Report Module. The user interface of this system can be used by user to organize the renewable energy generation reports daily, monthly, and yearly, to record the total renewable energy consumption (Watt), and to calculate the cost of net profit. Furthermore, after correcting the defects recorded during user acceptance test, the system can be effectively used for energy generation monitoring

work. This study contributes to the industry by showing how the IoT system can replace the manual data collection work for Hybrid Energy System. However, there are some limitations in this study that are recommended for further improvement. For example, there is currently no selection for graph types presented the homepage, either by daily, monthly or yearly records. Besides, the daily, monthly, and yearly savings should also be presented at the homepage module of software.

Acknowledgements This research was supported by Ministry of Higher Education (MOHE) through Fundamental Research Grant Scheme (FRGS) (FRGS/1/2020/TK0/UTHM/03/28) and the Universiti Teknologi Malaysia Encouragement Research (Grant No: PY/2020/04178).

References

1. P.P. Ray, A survey on internet of things architectures. *J. King Saud University-Comput. Inf. Sci.* **30**(3), 291–319 (2018)
2. Mottain, IR 4.0: On the brink of technological revolution. *The Star* (2019). Retrieved from: <https://www.thestar.com.my/business/business-news/2019/05/25/ir40-on-the-brink-of-technological-revolution>
3. M. Deepika, M. Kavitha, N.K. Chakravarthy, J.S. Rao, D.M. Reddy, B.M. Chandra, A critical study on campus energy monitoring system and role of IoT, in *2021 International Conference on Sustainable Energy and Future Electric Transportation (SEFET)* (IEEE, 2021), pp. 1–6
4. K. Jiju, P. Ramesh, P. Brijesh, B. Sreekumari, Development of Android based on-line monitoring and control system for renewable energy sources, in *2014 International Conference on Computer, Communications, and Control Technology (I4CT)* (IEEE, 2014), pp. 372–375
5. M.N.M. Shah, F. Ahmad, M.S. Abdullah, M.K. Musa, N.I. Abidin, H. Harun et al., Design and development of trash trap of stream for mini hydro, in *Materials Today: Proceedings* (2021)
6. W. Cole, N. Gates, T. Mai, Exploring the cost implications of increased renewable energy for the US power system. *Electricity J.* **34**(5), 106957 (2021)
7. J.J.D. Nesamalar, S. Suruthi, S.C. Raja, K. Tamilarasu, Techno-economic analysis of both on-grid and off-grid hybrid energy system with sensitivity analysis for an educational institution. *Energy Convers. Manage.* **239**, 114188 (2021)
8. B. Abdul Malek, Renewable energy development in Malaysia. Ministry of Energy, Green Technology and Water Malaysia (KeTTHA) (2011)