Pertanika Journal of

SCIENCE & TECHNOLOGY

JST

VOL. 27 (4) OCT. 2019

A scientific journal published by Universiti Putra Malaysia Press

About the Journal

Overview

Pertanika Journal of Science & Technology (JST) is the official journal of Universiti Putra Malaysia published by UPM Press. It is an open-access online scientific journal which is free of charge. It publishes the scientific outputs. It neither accepts nor commissions third party content.

Recognized internationally as the leading peer-reviewed interdisciplinary journal devoted to the publication of original papers, it serves as a forum for practical approaches to improving quality in issues pertaining to science and engineering and its related fields.

JST is a **quarterly** (January, April, July and October) periodical that considers for publication original articles as per its scope. The journal publishes in **English** and it is open to authors around the world regardless of the nationality.

The Journal is available world-wide.

Aims and scope

Pertanika Journal of Science and Technology aims to provide a forum for high quality research related to science and engineering research. Areas relevant to the scope of the journal include: bioinformatics, bioscience, biotechnology and bio-molecular sciences, chemistry, computer science, ecology, engineering, engineering design, environmental control and management, mathematics and statistics, medicine and health sciences, nanotechnology, physics, safety and emergency management, and related fields of study.

History

Pertanika was founded in 1978. A decision was made in 1992 to streamline Pertanika into three journals as Journal of Tropical Agricultural Science, Journal of Science & Technology, and Journal of Social Sciences & Humanities to meet the need for specialised journals in areas of study aligned with the interdisciplinary strengths of the university.

After almost 27 years, as an interdisciplinary Journal of Science & Technology, the revamped journal now focuses on research in science and engineering and its related fields.

Goal of *Pertanika*

Our goal is to bring the highest quality research to the widest possible audience.

Quality

We aim for excellence, sustained by a responsible and professional approach to journal publishing. Submissions are guaranteed to receive a decision within 14 weeks. The elapsed time from submission to publication for the articles averages 5-6 months.

Abstracting and indexing of *Pertanika*

The journal is indexed in SCOPUS (Elsevier), Clarivate-Emerging Sources Citation Index [ESCI (Web of Science)], BIOSIS, National Agricultural Science (NAL), Google Scholar, MyCite and ISC.

Future vision

We are continuously improving access to our journal archives, content, and research services. We have the drive to realise exciting new horizons that will benefit not only the academic community, but society itself.

Journal of Science & Technology
Journal of Science & Technology

Citing journal articles

The abbreviation for Pertanika Journal of Science & Technology is *Pertanika J. Sci. Technol.*

Publication policy

Pertanika policy prohibits an author from submitting the same manuscript for concurrent consideration by two or more publications. It prohibits as well publication of any manuscript that has already been published either in whole or substantial part elsewhere. It also does not permit publication of manuscript that has been published in full in Proceedings.

Code of Ethics

The Pertanika Journals and Universiti Putra Malaysia takes seriously the responsibility of all of its journal publications to reflect the highest in publication ethics. Thus all journals and journal editors are expected to abide by the Journal's codes of ethics. Refer to Pertanika's **Code of Ethics** for full details, or visit the Journal's web link at http://www.pertanika.upm.edu.my/code_of_ethics.php

International Standard Serial Number (ISSN)

An ISSN is an 8-digit code used to identify periodicals such as journals of all kinds and on all media–print and electronic. All Pertanika journals have ISSN as well as an e-ISSN.

Journal of Science & Technology: ISSN 0128-7680 (*Print*); ISSN 2231-8526 (*Online*).

Lag time

A decision on acceptance or rejection of a manuscript is reached in 3 to 4 months (average 14 weeks). The elapsed time from submission to publication for the articles averages 5-6 months.

Authorship

Authors are not permitted to add or remove any names from the authorship provided at the time of initial submission without the consent of the Journal's Chief Executive Editor.

Manuscript preparation

Refer to Pertanika's **INSTRUCTIONS TO AUTHORS** at the back of this journal.

Editorial process

Authors are notified with an acknowledgement containing a *Manuscript ID* on receipt of a manuscript, and upon the editorial decision regarding publication.

Pertanika follows a **double-blind peer-review** process. Manuscripts deemed suitable for publication are usually sent to reviewers. Authors are encouraged to suggest names of at least three potential reviewers at the time of submission of their manuscript to Pertanika, but the editors will make the final choice. The editors are not, however, bound by these suggestions.

Notification of the editorial decision is usually provided within ten to fourteen weeks from the receipt of manuscript. Publication of solicited manuscripts is not guaranteed. In most cases, manuscripts are accepted conditionally, pending an author's revision of the material.

The Journal's peer-review

In the peer-review process, three referees independently evaluate the scientific quality of the submitted manuscripts.

Peer reviewers are experts chosen by journal editors to provide written assessment of the **strengths** and **weaknesses** of written research, with the aim of improving the reporting of research and identifying the most appropriate and highest quality material for the journal.

Journal of Science & Technology
Journal of Science & Technology

Operating and review process

What happens to a manuscript once it is submitted to *Pertanika*? Typically, there are seven steps to the editorial review process:

1. The Journal's Chief Executive Editor (CEE) and the Editorial Board Members (EBMs) examine the paper to determine whether it is appropriate for the journal and should be reviewed. If not appropriate, the manuscript is rejected outright and the author is informed.

2. The CEE sends the article-identifying information having been removed, to three reviewers who are specialists in the subject matter represented by the article. The CEE requests them to complete the review in three weeks.

   Comments to authors are about the appropriateness and adequacy of the theoretical or conceptual framework, literature review, method, results and discussion, and conclusions. Reviewers often include suggestions for strengthening of the manuscript. Comments to the editor are in the nature of the significance of the work and its potential contribution to the field.

3. The CEE, in consultation with the Editor-in-Chief (EiC), examines the reviews and decides whether to reject the manuscript, invites the author(s) to revise and resubmit the manuscript. The CEE may seek additional reviews. Final acceptance or rejection rests with the CEE and EiC, who reserve the right to refuse any material for publication. In rare instances, the manuscript is accepted with almost no revision. Almost without exception, reviewers' comments (to the author) are forwarded to the author. If a revision is indicated, the editor provides guidelines for attending to the reviewers' suggestions and perhaps additional advice about revising the manuscript.

4. The authors decide whether and how to address the reviewers' comments and criticisms and the editor's concerns. The authors return a revised version of the paper to the chief executive editor along with specific information describing how they have answered' the concerns of the reviewers and the editor, usually in a tabular form. The author(s) may also submit a rebuttal if there is a need especially when the author disagrees with certain comments provided by reviewer(s).

5. The CEE sends the revised paper out for re-review. Typically, at least one of the original reviewers will be asked to examine the article.

6. When the reviewers have completed their work, the CEE in consultation with the EiC and EBMs examine their comments and decide whether the paper is ready to be published, needs another round of revisions, or should be rejected.

7. If the decision is to accept, an acceptance letter is sent to all the author(s), the paper is sent to the Press. The article should appear online in approximately three months.

   The Publisher ensures that the paper adheres to the correct style (in-text citations, the reference list, and tables are typical areas of concern, clarity, and grammar). The authors are asked to respond to any minor queries by the Publisher. Following these corrections, page proofs are mailed to the corresponding authors for their final approval. At this point, **only essential changes are accepted**. Finally, the article appears in the pages of the Journal and is posted online.

Pertanika Journal of

SCIENCE & TECHNOLOGY

Vol. 27 (4) Oct. 2019

PERTANIKA
JOURNALS

A scientific journal published by Universiti Putra Malaysia Press

JST
Journal of Science & Technology
AN INTERNATIONAL PEER-REVIEWED JOURNAL

EDITORIAL OFFICE

JOURNAL DIVISION
Office of the Deputy Vice Chancellor (R&I)
1st Floor, IDEA Tower II
UPM-MTDC Technology Centre
Universiti Putra Malaysia
43400 Serdang, Selangor Malaysia.
Gen Enq.: +603 9769 1622 | 1616
E-mail:
executive_editor.pertanika@upm.edu.my
URL: www.journals-jd.upm.edu.my

PUBLISHER
UPM Press
Universiti Putra Malaysia
43400 UPM, Serdang, Selangor, Malaysia.
Tel: +603 9769 8851
E-mail: penerbit@putra.upm.edu.my
URL: http://penerbit.upm.edu.my

ABSTRACTING AND INDEXING OF PERTANIKA JOURNALS

Pertanika has reached **40 years old**; this accumulated knowledge has resulted in the journals being abstracted and indexed in SCOPUS (Elsevier), Clarivate-Emerging Sources Citation Index [ESCI (Web of Science)], BIOSIS, National Agricultural Science (NAL), Google Scholar, MyCite and ISC.

PERTANIKA
JOURNALS

PENERBIT UPM
UNIVERSITI PUTRA MALAYSIA
PRESS

Pertanika Journal of Science & Technology
Vol. 27 (4) Oct. 2019

Contents

Foreword — i
Abu Bakar Salleh

Medical and Health Sciences

Review article
Medical Imaging Literature in MyCite — 1471
Nurul Husna Kamarudin, Nor Azlina A. Rahman and Zainul Ibrahim Zainuddin

Review article
Psychosocial Predictors of Adolescent Aggression — 1485
Fatimah Ahmad Fauzi, Nor Afiah Mohd Zulkefli, Halimatus Sakdiah Minhat and Norliza Ahmad

Review article
Timing and Prognostic Factors of Tuberculosis Treatment Interruption — 1509
Qudsiah Suliman, Salmiah Md. Said and Lim Poh Ying

Simulation Work on Blood Glucose Control for Type 1 Diabetes using Modified Hovorka Equations — 1527
Ayub Md Som, Nur Farhana Mohd Yusof, Sherif Abdulbari Ali and Nurafiqah Zawawi

Composite Resin Veneer Systems: An *In Vitro* Study to Evaluate Surface Roughness Changes with Different Surface Treatments and Shear Bond Strength — 1539
Nik Zarina Nik Mahmood, Amar Mohammed Thiyab and Mohamed Ibrahim Abu Hassan

Exploring Potential Neuroprotective Properties of Aqueous *Centella asiatica* Extract in Chronic Stress-induced Rats — 1561
Muhamad Zulhusni Abdul Wahab, Muhammad Zaim Zainal Abidin, Fatin Nadzirah Zakaria, Mohamad Aris Mohd Moklas, Zulkhairi Amom, Nina Keterina Hashim and Razif Dasiman

Forensic Parameter and Paternity Testing of Tpa-25 Element in Kelantan-Malay and Jawa-Malay — 1575
Nurul Syafika Muslimin, Nur Hilwani Ismail and Farida Zuraina Mohd Yusof

Effect of Diet Regime on the Development and Survival of *Aedes albopictus* (Skuse) (Diptera: Culicidae) — 1589
Nurul Syafika Muslimin, Nur Hilwani Ismail and Farida Zuraina Mohd Yusof

Monitoring of Fracture Healing Process by Acousto-Ultrasonic Technique 1603
Sharad Shrivastava and Ravi Prakash

MRN Complex and ATM Kinase Inhibitors Impacts towards UVC-Treated Zebrafish Embryonic Development 1615
Shazrul Fazry, Malina Kumaran, Nahid Khalili, Ibrahim Mahmood, Herryawan Ryadi Eziwar Dyari, Nur Hidayah Jamar, Douglas Law, Nik Marzuki Sidik, Nor Azian Abdul Murad and Yi Chieh Lim

Identifying Analogues Of 2-Deoxyglucose, Alpha-D-Glucose and Beta-D-Glucose-6-Phosphate as Potential Inhibitors of Human Hexokinase II for the Development of Anti-Dengue Therapeutics 1625
Suriyea Tanbin, Nurhainis Ogu Salim and Fazia Adyani Ahmad Fuad

Information, Computer and Communication Technologies

Review article
Feature Extraction and Classification Techniques of MODI Script Character Recognition 1649
Solley Joseph and Jossy George

Resolution Adaptive Threshold Selection for Gradient Edge Predictor in Lossless Biomedical Image Compression 1671
Urvashi, Meenakshi Sood and Emjee Puthooran

Hybrid Feature Extraction and Machine Learning Approach for Fruits and Vegetable Classification 1693
Nimratveer Kaur Bahia, Rajneesh Rani, Aman Kamboj and Deepti Kakkar

A New Method to Construct 4-Cycle-Free Parity-Check Matrices for Regular LDPC Codes 1709
Djamel Slimani, Adda Ali-Pacha and Naima Hadj-Said

Matching Fingerprint Images for Biometric Authentication using Convolutional Neural Networks 1723
Abdulmawla Najih, Syed Abdul Rahman Al-Haddad Syed Mohamed, Abdul Rahman Ramli Shaiful Jahari Hashim and Nabila Albannai

Improved Salient Object Detection via Boundary Components Affinity 1735
Nur Zulaikhah Nadzri, Mohammad Hamiruce Marhaban, Siti Anom Ahmad, Asnor Juraiza Ishak and Zalhan Mohd Zin

Detection and Recognition via Adaptive Binarization and Fuzzy Clustering 1759
Saad Mohmad Saad Ismail, Siti Norul Huda Sheikh Abdullah and Fariza Fauzi

Time Adaptive Collaborative Filtering for Movie Recommendation 1783
Suganeshwari Gopalswamy and Syed Ibrahim Peer Mohamed

Criminality Recognition using Machine Learning on Malay Language Tweets 1803
Nurul Hashimah Ahamed Hassain Malim, Saravanan Sagadevan and Nurul Izzati Ridzuwan

Measuring Tajweed Augmented Reality-Based Gamification Learning Model (TARGaLM) Implementation for Children in Tajweed Learning 1821
Nurtihah Mohamed Noor, Marina Ismail, Rahmah Lob Yussof and Fakhrul Hazman Yusoff

Environmental Sciences

Review article
Sago Wastes and Its Applications 1841
Nurhusni Amin, Nordin Sabli, Shamsul Izhar and Hiroyuki Yoshida

Optimization of Process parameters for Decolorization of Azo Dye Remazol Golden Yellow by *Bacillus firmus* using Biostatistical Designs 1863
Palanivelan Ramachandran, Subashni Bhoopathy, Ramya Suseenthar and Ayyasamy Pudukkadu Munusamy

Water Quality Effects on the Optimal Water Resources Operation in Great Karun River Basin 1881
Iman Ebrahim Bakhsipoor, Seyed Mohammad Ashrafi and Arash Adib

Evaluation of Aerosol Optical Thickness over Malaysia Based on Multi-Source Ground and Satellite Data 1901
Khaled Ali Ahmed Ben Youssef, Ahmad Makmom Abdullah, Helmi Zuhaidi Mohd Shafri and Zulfa Hanan Ash'aari

Growth Performance of Blood Cockle (Tegillarca granosa) within Kongkong Laut Estuaries, Masai, Johor 1917
Amirul Azuan Md Joni, Ferdaus Mohamat Yusuff, Khairul Nizam Mohamed, Faradiella Mohd Kusin and Syaizwan Zahmir Zulkifli

Spatial and Vertical Metals Distribution in Sediment Cores from Kongkong Laut Estuary, Masai, Johor 1929
Sara Azuana Khalidah Khalid, Ferdaus Mohamat Yusuff, Faradiella Mohd Kusin, Aziz Arshad and Zulfa Hanan Asha'ari

Concentration of Heavy Metals in Street Dust and Surface Soils in Urban and Peri-Urban Regions in the Kuala Lumpur Metropolitan Area 1941
Ahmad Ismail, Munirah Hanapiah, Syaizwan Zahmir Zulkifli and Hishamuddin Omar and Nur Amiera Kamarudin

Long-Term Water Quality Assessment in a Tropical Monsoon 1955
Jabir Harna Abdulkareem, Da'u Abba Umar, Alhassan Idris Gabasawa, Chinedum Anyika and Nor Rohaizah Jamil

Case study
Monitoring Urban Green Space (UGS) Changes by Using High Resolution Aerial Imagery: A Case Study of Kuala Lumpur, Malaysia 1971
Junainah Abu Kasim, Mohd Johari Mohd Yusof and Helmi Zulhaidi Mohd Shafri

Engineering Sciences

Computational Study of Mixed Heat Convection in Annular Space between Concentric Rotating Inner and Wavy Surface Outer Cylinders 1991
Qusay Rasheed Al-Amir, Farooq Hassan Ali Alinnawi and Qusay Adnan

Strategies to Improve Communication Management within Virtual Project Teams 2015
Nuzul Azam Haron, Law Teik Hua, Salihudin Hassim, Fathollah Eftekhari, Muhammad Tahir Muhammad and Aizul Nahar Harun

Assessment of External Risk Factors on Construction Project Schedule Using Risk Importance Index (RII) 2031
Saiful Husin, Abdullah, Medyan Riza and Mochammad Afifuddin

Detection of Nitrobenzene using a coated Quartz Crystal Microbalance with a Parametric Modeling Approach 2049
Dipali Ramdasi and Rohini Mudhalwadkar

Comparison of Overall Heat Transfer Coefficient between Shell and Tube and Spiral Coil Heat Exchangers 2065
Meseret Nasir Reshid, Girma Tadesse Chala and Wan Mansor Wan Muhamad

Optimization of Glass-beads Water Slurry flow Characteristics using Taguchi's Decision-making Technique coupled with Genetic Algorithm 2079
Shofique Uddin Ahmed, Rajesh Arora and Ranjana Arora

Case study
An Eco-Innovative Green Design Method by QFD and TRIZ Tools - A Case Study of Brass-Ware Manufacturing 2109
Anhad Malhotra, Sonu Rajak and Sanjay Kumar Jha

Utilisation of *Cassia surattensis* Seeds as Natural Adsorbent for Oil Content Removal in Oilfield Produced Water 2123
Mohmad Khalel Ibrahim, Ali Abd Al-Hassan and Ahmed Samir Naje

Applied Sciences and Technologies

Modelling and Optimization of Torrefied Pellet Fuel Production — 2139
Ahmad Hafizi Awang, Abdulhalim Abdulrazik, Azuin Mad Noor and Aainaa Izyan Nafsun

Techniques and Algorithms for Selection of Number and Locations of Temperature Sensors for Greenhouse — 2153
Suman Lata and Harish Kumar Verma

Three Phase Grid Connected Neutral Point Clamped (NPC) Multilevel Inverter Fed by Two Wind Turbines — 2173
Sanusi Alabi Kamilu and Saad Mekhilef

Efficacy of Silvernano Bacterial Cellulose Insole on *Micrococcus sedentarius* Growth — 2189
Kanchana Sitlaothaworn and Manussawee Dechkla

Elemental Concentration in Three Different Fish Species Captured from Oluwa River, Okitipupa, Ondo State, Nigeria — 2201
Thompson Faraday Ediagbonya and Ramond Babatunde Adesokan

Fatty Acid Profile and Production of Fatliquor from *Canarium schweinfurthii* Mesocarp Oil — 2221
Adachukwu Nwaka Nkwor, Pius Oziri Ukoha, William Robert Wise, Njemuwa Njoku Nwaji and Karl Flowers

Earth Sciences

Assessment of Runoff via Precipitation using Neural Networks: Watershed Modelling for Developing Environment in Arid Region — 2245
Sandeep Samantaray, Abinash Sahoo and Dillip Kumar Ghose

Prediction of Groundwater Contaminants from Cattle Farm using Visual MODFLOW — 2265
Mohammad Nazri Ebrahim, Hasfalina Che Man, Mohamed Azwan Mohamed Zawawi and Muhammad Hazwan Hamzah

Spatial Pattern and Temporal Variability of Sea Level Anomaly and Geostrophic Current in the Eastern Indian Ocean from Satellite Altimetry — 2281
Julian Saputra, Jonson Lumban-Gaol, James Parlindungan Panjaitan and Agus Saleh Atmadipoera

Strength Assessment of Mining Soil Treated with Steel Slag as Liner Material for Retention Pond — 2305
Noratiqah Masri, Faradiella Mohd Kusin, Nik Norsyahariati Nik Daud and Hasfalina Che Man

Distribution of Dissolved and Particulate Fe in an Estuarine System at Bagan Pasir, Perak — 2317
Syaza Syafiqah Muhamad Kamel, Khairul Nizam Mohamed, Ferdius@Ferdaus Mohamat Yusuff and Abd Muhaimin Amiruddin

Hydrological Response of Semi-Arid River Catchment to Rainfall and Temperature Fluctuations 2333
Da'u Abba Umar, Mohammad Firuz Ramli, Ahmad Zaharin Aris and Muhammad Amar Zaudi

Mathematical Sciences

Some Relation Properties of Rough Neutrosophic Multisets 2351
Suriana Alias, Daud Mohamad and Adibah Shuib

Short communication
Characterization of Model on Zoning Process 2373
Ompal Singh

Two Point Block Multistep Methods with Trigonometric–Fitting for Solving Oscillatory Problems 2381
Aini Fadhlina Mansor, Fudziah Ismail and Norazak Senu

Chemical Sciences

Effect of Parboiling on *in vitro* Physiological Antioxidant Capacity of Brown Rice 2399
Elaine Lee, Phene Neoh Pei Nee, Lee Huei Hong, Wong Sie Chuong, Wong Tze Jin, Koo Lee Feng and Yiu Pang Hung

Characteristics of Chitosan Nanoparticles extracted from Sea Cucumber (*Holothuria scabra*) as Source Materials for Glucosamine 2409
Sumarto, Bustari Hasan, Rahman Karnila and Mery Sukmiwati

Material Sciences

Preparation and Characterization of Zeolite type 4A using Kaolin from Ajebo, Nigeria 2427
Henry Ekene Mgbemere, Ikenna Christopher Ekpe, Ganiyu Lawal, Henry Ovri and Anna-Lisa Chaudhary

Physical Sciences

Scattered Radiation Dose Calculation in the Multi-Bay Resuscitation Room using MCNP5 Code 2439
Marianie Musarudin, Ruzainah Mohd Nordin and Nik Kamarullah Ya Ali

Foreword

Welcome to the fourth Issue of 2019 for the Journal of Science and Technology (JST)!

JST is an open-access journal for studies in Science and Technology published by Universiti Putra Malaysia Press. It is independently owned and managed by the university for the benefit of the world-wide science community.

This issue contains 57 articles; 5 are review articles, 1 case study, 1 short communication and the rest are regular articles. The authors of these articles come from different countries namely Malaysia, United Kingdom, Germany, Denmark, Indonesia, Nigeria, India, Oman, Algeria, Iraq, Iran and Thailand.

Nik Zarina Nik Mahmood and colleagues from Universiti Teknologi MARA discussed about the surface roughness and shear bond strength of 2 types of composite veneer. They concluded that the laboratory-made veneer system exhibited superior surface roughness than prefabricated veneer system when treated with different types of surface treatments. Further details on the study can be found on page 1539.

A regular article titled "Identifying analogues of 2-deoxyglucose, α-D-glucose and β-D-glucose-6-phosphate as potential inhibitors of human hexokinase II for the development of anti-dengue therapeutics" was co-authored by Suriyea Tanbin, Nurhainis Ogu Salim and Fazia Adyani Ahmad Fuad from International Islamic University Malaysia and Institute for Medical Research. They successfully identified compounds similar to 2-deoxyglucose, α-D-glucose and β-D-glucose-6-phosphate by using ligand-based screening program. Detailed information on this study can be found on page 1625.

Mohmad Khalel Ibrahim and researchers from University of Al Muthanna and Al-Qasim Green University, Iraq summarized the utilization of *Cassia surratensis* seeds as a natural adsorbent for oil content removal in oilfield produced water. They concluded that the use of Cassia surattensis seeds to eliminate oil content in oilfield produced water was an ecological and low-cost method, which was confirmed with the Freundlich adsorption isotherm model. Details on this study are presented on page 2123.

We anticipate that you will find the evidences presented in this issue to be intriguing, thought provoking and useful in reaching new milestones in your own research. Please recommend the journal to your colleagues and students to make this endeavour meaningful. All the papers published in this edition underwent Pertanika's stringent peer-review process involving a minimum of two reviewers comprising internal as well as external referees. This was to ensure that the quality of the papers justified the high ranking of the journal, which is renowned as a heavily-cited journal not only by authors and researchers in Malaysia but by those in other countries around the world as well.

We would also like to express our gratitude to all the contributors, namely the authors, reviewers, Editor-in-Chief and Editorial Board Members of JSSH who have made this issue possible.

JST is currently accepting manuscripts for upcoming issues based on original qualitative or quantitative research that opens new areas of inquiry and investigation.

Chief Executive Editor
Prof. Dato' Dr. Abu Bakar Salleh
executive_editor.pertanika@upm.edu.my

Pertanika J. Sci. & Technol. 27 (4): 1471 - 1483 (2019)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Review article

Medical Imaging Literature in MyCite

Nurul Husna Kamarudin[1], Nor Azlina A. Rahman[2] and Zainul Ibrahim Zainuddin[1]*

[1]*Department of Diagnostic Imaging and Radiotherapy, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, Kuantan, Pahang, Malaysia*
[2]*Department of Physical Rehabilitation Sciences, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, Kuantan Pahang, Malaysia*

ABSTRACT

The presence of imaging technologies in Malaysia needs to be supported by homegrown research to optimize and tailor their usage for local benefits. Research done elsewhere may not be applicable to local situations. This study investigates the contribution of researches by Malaysian academicians and service providers to the field of medical imaging, as evident in the Malaysian Citation index (MyCite) database. Bibliometric and thematic analyses were performed on publications featured in the database from 2006 to 2016. The bibliometric analysis provided information on the affiliation of the authors, their professional backgrounds, types of studies, and the journals involved while the thematic analysis identified the themes and sub-themes of identified articles. The study found that Malaysians contributed 54.1% of the publications, followed by non-Malaysians (41.8%) and collaboration authors (4.1%). Researchers were mostly from university-based and hospital university-based institutions. The thematic analysis presented that 42.9% of articles were classified under clinical themes. The results also suggested that the current trends of research in medical imaging are focused on topics involving clinical and modality, and only a few patient-centered researches. This is an indication that more researches that are relevant to local practices and needs are required as this will strengthen the medical imaging practice in the country.

Keywords: Academicians, bibliometric study, medical imaging, MyCite, Scopus, service providers, thematic analysis, web of science

ARTICLE INFO

Article history:
Received: 06 January 2018
Accepted: 23 April 2019
Published: 21 October 2019

E-mail addresses:
husnataqwa@gmail.com (Nurul Husna Kamarudin)
nazara71@gmail.com (Nor Azlina A. Rahman)
zainul@iium.edu.my (Zainul Ibrahim Zainuddin)
* Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

INTRODUCTION

Recent technological developments in medical imaging underline the significance of research in the profession. Research is needed to provide the evidence for its practices as research is a means by which a profession defines its own knowledge base and successfully differentiates itself from others (Reeves, 2008). The importance of research and the dissemination of the findings could facilitate the growth of the discipline. Research in medical imaging is further dictated by the concept of evidence-based practice (EBP), a concept that is aimed to enable healthcare practitioners to deliver optimal healthcare services. It is important for medical imaging practitioners to embrace this concept to remain relevant in the healthcare industry.

Malaysia is a rapidly progressing nation that aims to be a fully developed country by the year 2020. The country embarks on efforts to strengthen the industry and research collaboration in order to achieve a fully developed country standard (Economic Planning Unit, 2010). While those aspirations have been formulated, an important undertaking to determine whether the current status of locally grown knowledge and researches are in line with the aspirations of the nation. Homegrown research is significant as researchers will conduct their research based on the needs of the nation. Research conducted elsewhere may not be directly applicable to the local needs due to sociobiological, socioeconomic, cultural and religious factors. Thus, it is felt that the researchers need to support the government's outlook in prioritizing their research interests and undertakings.

Scientific writings through journal publications are the more commonly accepted mode of disseminating research findings. In 2011, the Ministry of Education Malaysia (MOE) established the Malaysian Citation Centre (MCC) that serves to collate, monitor, coordinate and improve the standards of journal publications in Malaysia (MyCite, 2017). This was followed by the setting up of the Malaysian Citation Index, MyCite. This online database provides access to bibliographic, as well full-text contents of scholarly journals, conference proceedings, book chapters and theses in science, technology, medicine, social science and humanities published in Malaysia. To date, it has more than 34,000 documents. The MyCite Journal Selection criteria have to be met before any journal can be considered for a MyCite indexation. An important selection criterion for MyCite database is international diversity, indicated by 40% international authors and 60% local (regional) authors. Priority is also given to a journal if it has achieved a reasonable degree of influence as a result of its indexation status in Web of Science (WoS) or Scopus, or any discipline-based databases.

The trends and priorities of research play an important role to reap the optimum benefits to the service and education of any given profession. Thus, the importance of research in medical imaging field among its academicians and service providers cannot be sidelined. It would be beneficial to determine how these researches are seen within the context of meeting the expectations of the overall medical imaging field in Malaysia. It is felt that

MyCite is a significant database to be investigated in order to determine the current research trends in medical imaging in Malaysia. Thus, the objective of this study is to examine the trend of publications in medical imaging as indexed in the database. This study is hoped to convince future researchers to acknowledge the trends of research in medical imaging in Malaysia and guide them towards making informed decisions about future researches in the discipline.

MATERIALS AND METHODS

There are ample literatures that suggest the use of bibliometric method as a tool in the assessment of research initiatives and performances in basic scientific disciplines and medicine. The bibliometric method is a qualitative study of publication that is used to describe the patterns of publications within the given field (Clarke et al., 2007). The method is also useful to map the literature in a given field by identifying patterns of authorship and productivity in providing a direction for future development of the respective field. It can also provide an in-depth analysis of the connections between authors, publications and research topics (Snaith, 2012).

Several medical imaging related terms were used as search terms in this study. These included radiography, medical imaging, radiology, radiation, X-ray, neuroimaging, ultrasound, magnetic resonance imaging (MRI), computed tomography (CT) scan, nuclear imaging, mammography, intravenous urogram (IVU), and fluoroscopy. The researchers caution that these terms are not exhaustive, yet they were felt to be sufficient for the purpose of the study.

The MyCite database (http://www.mycite.my) was accessed. The Advanced Search tab was chosen to accommodate more than one search term. Using the Boolean operator "OR" and additional search tabs, the medical imaging terms mentioned above were individually entered into the search tabs. A general survey involving 'All Field' was done. The search was limited to articles and published in English and Malay language. The search covered a 10-year period from 2006 to 2016. This time frame was chosen to depict the recent publications that complemented the changes experienced by the field of medical imaging and was not restricted towards the authors' countries of origin. The total number of articles that resulted from the Advanced Search was noted. Information that included the authors' names, titles of articles, journal titles, year of publication and abstracts were manually tabulated in an Excel file.

The titles of the articles were arranged alphabetically. The titles of articles were checked for duplication before it was analysed by researchers. The "filter" tab in the Excel file was used to recognize those titles and duplications were removed from the data. The filtered data was further focused to articles related to medical imaging. Since the titles of the articles represented the main idea, they were used as indicators to determine whether the articles were related to medical imaging or otherwise. Titles that were not related Medical imaging such as computer science, physics and astronomy were excluded from the data. There were some titles that did not represent medical imaging. But, they were included after reviewing the abstract of the articles due to the issues discussed in the articles were strongly related to medical imaging. The final data was used for bibliometric and thematic studies.

Fulfilling the characteristics of a bibliometric study, the affiliations of the authors, types of studies, and professional background of the authors was included in this study. The affiliations determined whether the articles were contributed by authors from Malaysia, or collaborations between Malaysian and non-Malaysian authors. The data was also used to classify the publications into review, comparative study, survey and empirical study. Special emphasis had been given to publications by Malaysian researchers. The determination of the background of the authors served to indicate whether the authors were from hospitals, universities, hospital universities, government bodies, private bodies or non-governmental organizations (NGOs). Further classification divided the authors into those who were from the clinical, academia or industry related to medical imaging. For the purpose of this study, a clinical provider is one who works in a clinical field in medical imaging, whereas an academician relates to an academic staff at a higher educational institution in Malaysia. Meanwhile, an industry worker is a person who works with any industrial company that is related to the medical imaging field.

Thematic analysis, a part of qualitative research is used to identify, analyse, and report patterns (themes) within the data (Braun & Clarke, 2006). The analysis provides a robust, systematic framework for coding qualitative data and for using the coding to define patterns across the dataset in researches. It is a pragmatic qualitative approach for those doing applied research and has been used widely for health and well-being research (Braun & Clarke, 2014). For the present study, thematic analysis was used to define the significant themes and sub-themes based on the titles of articles. This was done for the themes denoted the major issues being discussed, while the sub-themes denoted the related themes to the main issues within the titles.

RESULTS AND DISCUSSIONS

The General Report

Data was collected on 29th November 2016. Although there were 32,926 documents indexed in MyCite, the number of journals that fulfilled the MyCite Journal Selection Criteria stood at 138. MyCite estimated that more than 500 Malaysian journals can benefit from being indexed in the database in terms of global visibility and enabling Malaysian researchers to identify expertise, areas of possible collaboration, stimulate use and citations (MyCite, 2017). These estimation is credible since it was presented in 2012 that there were 424 traceable journal titles that were published in Malaysia (Zainab et al., 2012). In the same work, the authors reported a finding in 1999 that there were 284 titles. This means there was a 52% increase in journal titles within 13 years. Thus, with only 138 journals indexed in MyCite, from a possible of 500 Malaysian journals, the 27% indexation rate in the database shows much support is needed to ensure more journals to be cited in the database. With MyCite providing access to bibliographic as well full-text contents of scholarly journals published in Malaysia in the various fields, practitioners should optimise the presence of the database. Primarily, "homegrown" research, translated into scientific writings indexed in MyCite can be benefitted by other local practitioners since MyCite provides the necessary access to those researches. Thus, the two important issues; support for local journals to be indexed in MyCite and the accessibility of publications in the database could have their impact upon practitioners in all fields, including medical imaging, in managing their research and disseminating the findings for local practitioners.

Report for Medical Imaging Literature

The study found the number of articles that were related to medical imaging in MyCite was 268. This represents 0.8% of the total number of articles. 156 articles were contributed by Malaysians while 112 were from non-Malaysians. The results show that the 268 articles were published in 32 journals. This represents 23% of the total number of journals indexed. The top 5 journals that published articles related to medical imaging are listed in Table 1.

Only 40% of the journals indexed in MyCite addressed the science, technology and medical fields. Being a sub-discipline in the Medicine field, the low number of articles related to medical imaging is expected. However, this study does not represent an overall true picture of the contributions of Malaysian medical imaging practitioners to their field in terms of research. It is believed that there are articles by Malaysian practitioners being indexed in some other databases such as Scopus, WoS or even in other non-indexed journals. The lower output of researches related to medical imaging in MyCite could be attributed to the inclination of researchers to publish their work in journals indexed in Scopus, WoS or other databases. However, the accessibility to those mentioned databases may be limited and subjected to subscriptions requirements as it is often confined to academic institutions. Thus, it is important to reiterate that "homegrown" researches should be made readily

Table 1
Top five journals that published articles related to medical imaging

No.	Name of Journal	Total articles (%)
1.	Biomedical Imaging Intervention Journal	107 (40%)
2.	Medical Journal of Malaysia	27 (10%)
3.	Malaysian Journal of Medical Science	27 (10%)
4.	Jurnal Teknologi	16 (6%)
5.	Neurology Asia	10 (4%)

accessible to local practitioners and is provided by MyCite. It is to be appreciated that there are nearly 25% of the journals in MyCite that had published articles related to medical imaging. This is translated into available avenues and opportunities for Malaysian medical imaging practitioners to publish their findings for the benefits of other local practitioners.

The study also identified 15 out of 32 journals being indexed in Scopus, while 4 out of 15 journals were indexed in WoS. They are shown in Table 2.

The results should be seen within the context of dispelling the conflict among academicians, especially in publishing their works in high impact journals in databases compared to Malaysian-based journals. It is a common knowledge that academicians are encouraged to publish in journals in the above databases for personal promotion and self-development, as well as for the rating and ranking of the university. There is a hidden element of “patriotism” if one were to consider publishing in any one of the 15 journals given in Table 2. This element should be seen within the context of supporting local journals to remain in the two mentioned databases. This support is needed for those journals to fulfil the criteria set by these databases in order to preserve the indexation. By preserving the indexation by the said journals in WoS and Scopus, the opportunity to further publish in journals indexed in the latter databases is simultaneous achieved.

Table 2
List of Journals that are indexed in Scopus and Web of Science

No.	Name of Journals	Scopus	Web of Science
1.	Biomedical Imaging Intervention Journal	√	
2.	International Medical Journal Malaysia	√	
3.	Journal of Mechanical Engineering	√	
4.	Malaysian Family Physician	√	
5.	Malaysian Journal of Analytical Sciences	√	
6.	Malaysian Journal of Computer Science	√	√
7.	Malaysian Journal of Mathematical Science	√	
8.	Malaysian Journal of Medical Sciences	√	
9.	Malaysian Journal of Medicine and Health Sciences	√	
10.	Malaysian Journal of Nutrition	√	
11.	Malaysian Journal of Public Health Medicine	√	
12.	Medical Journal of Malaysia	√	
13.	Neurology Asia	√	√
14.	Pertanika Journal of Science and Technology	√	√
15.	Sains Malaysiana	√	√

Country of Origin of Authors

The articles were classified into three groups: Malaysian authors, non-Malaysian authors and collaboration works. The total number of articles that were contributed by Malaysian authors was 145 articles (54.1%). Non-Malaysian authors contributed 112 articles (41.8%), while collaborative works between Malaysian and non-Malaysian authors were represented by 11 articles (4.1%). It can be expected that the total number of articles contributed by Malaysian authors to be higher than non-Malaysian authors as MyCite is a Malaysian online database. Thus, the contributions of Malaysian researchers to these journals are expected.

The national aspirations within the Malaysian academic environment lies in Malaysian authors are encouraged to publish in the higher impact journals. On a positive note, the presence of non-Malaysian authors in Malaysian scholarly journals is by itself a recognition to these local journals. Collaborations that are evident between Malaysian and non-Malaysian authors further enhance the position of the journals in MyCite. These collaborations can also increase the chances of these journals to be indexed in either WoS or Scopus later as the selection criteria are similar.

Affiliations of Malaysian Authors

The results show that affiliations of the authors are from universities (60), university hospitals (19), hospitals (5), government bodies (4) and private body (1). Contributions from academicians are prominent. This is expected since academicians are expected to not only publish, but also to simultaneously publish in high quality journals (Suryani et al., 2013). The affiliations of those representing university hospitals could still be from those within the academic circle, thus the expectations upon them to publish remains. An important note is the low numbers of authors from the hospitals. The term hospitals here refer to the hospitals that are not affiliated to any particular institute of higher learning. They are represented by the practitioners who are service providers. For 2016, Malaysia has about 140 public hospitals and 209 private hospitals (Ethnographic Medical Research Group, 2016). With only 5 authors representing hospitals, the results triggered further work, being undertaken by the present authors, to determine the reasons behind the low numbers.

Collaboration

A number of collaboration works by the authors were recognized. There were collaborations between universities (20), universities and government bodies (9), universities and hospital universities (7), universities and hospitals (7), and hospitals and hospital universities (3). Collaborations between universities and hospitals could be seen in areas where the hospitals provide the modalities or research subjects that are not available in the universities. These can be the case when clinical studies are involved. An example of a collaboration study between a university and hospital is the determination of the role of computed tomography (CT) scan in the assessment of parametrial involvement in early stage cervical carcinoma (Mohamad et al., 2012).

The study also took note of collaborations between local universities with non-local universities. For example, a collaborative study conducted by University of Nottingham, Malaysia campus with University of Cairo on brain magnetic resonance image lateral ventricles deformation analysis and tumour prediction (Kai et al., 2007). Though the number is small, there were also collaborations between universities with government bodies, private bodies or Non-Governmental Organisations (NGOs). Private body is an all for-profit business and not operated and funded by any governmental body. NGO refers to a non-profit organization which is independent from states and international governmental organization. The NGO is usually funded by donations and run primarily by volunteers. Universiti Kebangsaan Malaysia (UKM), Hospital Tengku Ampuan Rahimah and an NGO, The National Cancer Society of Malaysia conducted a study that determined tube output (kVp) and exposure mode for breast phantom of various thickness/glandurity for digital mammography (Kamal et al., 2015).

The results show that there are about 10 collaboration works involving hospitals with universities or university hospitals. This suggests that the contributions of those in the service are still low. With more than 300 hospitals several questions regarding the involvement of those in the service of medical imaging research can be raised. This can further be argued based on the fact that medical imaging graduates are now employed in both government and private hospitals. The authors wish to highlight that the education on a degree level in Medical imaging in Malaysia began in 2000 at UKM. This was followed by other public universities, namely Universiti Teknologi MARA (UiTM), International Islamic University Malaysia (IIUM) and two private universities; MAHSA University College and Masterskills. Top-up programmes from diploma to degree qualifications have also been introduced and graduates from these programmes have been produced. Modules involving research were introduced to these students, where the students were to conduct simple research to fulfil graduating criteria. Taking into consideration the number of students that have graduated since 2004, the number of researches that are translated into publications is expected to increase. The position of those who had graduated, in relation to the expected applications of research skills, will need to be examined in terms of their interest, as well as obstacles. Work is being undertaken by the authors to determine the involvement of these graduates in post-graduation research initiatives.

Thematic Analysis

The articles were categorized based on their types of studies. The result were tabulated in the Table 3.

A review helps authors to review published literature or data (Philip, 2009), thus establishing recent progress in a particular field. An example of a review is Bone health status and lipid profile among post-menopausal Malay women in Cheras, Kuala Lumpur by UKM and Universiti Putra Malaysia (UPM) (Hasnah et al., 2012). Meanwhile, an empirical study reports the results of a study that derived data from actual observation

Table 3
Thematic analysis of types of studies in MyCite

Type of study	Total (%)
Review studies	175 (65.2%)
Empirical studies	34 (12.7%)
Case studies	30 (11.2%)
Comparative studies	19 (7%)
Surveys	9 (3.4%)
Others	1 (0.4%)

or measured phenomena from experimentation (Bruns, 2010). This type of study helps academicians and clinical services to corporate the new findings into practices in medical imaging field. One such study is the application of Computer Tomographic (CT) data and additive manufacturing technologies in Prosthetic ear reconstruction involving SIRIM Berhad Malaysia and Universiti Sains Malaysia (USM) (Nor & Zainul, 2015). Surveys are used to compile data on a wide range of issues for different units of analysis (Rose et al., 2015). The main purpose of the survey research was to obtain information that describes the characteristics of a large sample of individuals of interest relatively (Ponto, 2015). A survey on physical factors and compressed breast thickness in voluntary mammography screening using FFDM system in Malaysia involving Malaysian Nuclear Agency, College of Radiology and IIUM is one of the examples for survey study as indexed in MyCite database (Noriah et al., 2013).

The thematic analysis for each title of the articles concentrated on the themes and sub-themes. The five most significant themes and subthemes were tabulated in Table 4.

The clinical-diagnosis, modality-technical and safety-radiation safety being the most significant themes and sub-themes for Malaysian authors. The clinical-diagnosis denoted that the function of medical imaging is to rule out any clinical disease. An example of a research title with the theme of clinical-diagnosis is Computed Tomography Perfusion Imaging on Traumatic Cerebral Contusion: A Preliminary Report by Universiti Sains Malaysia (USM) (Abdul et al., 2010). The clinical-health technology assessment represents the research in medical imaging that is involved in innovation in technology to improve the

Table 4
Top five themes and subthemes for thematic analysis

Themes	Subthemes	Total (%)
Clinical	Diagnosis	59 (22%)
Clinical	Health Technology Assessment	28 (10.5%)
Modality	Technical	27 (10%)
Safety	Radiation safety	24 (9%)
Clinical	Modality	14 (5.2%)

effectiveness in clinical area such as combined ultrasound and IVU for the management of childhood urolithiasis: a case report by Fathinul and Abdul (2012). Meanwhile, the modality-technical is a research involved in technical aspects on modalities in medical imaging. An example of a research in modality-technical is Tomotherapy as a Tool in Image-guided Radiation Therapy (IGRT): Theoretical and Technological Aspects (Yartsev et al., 2007). For safety-radiation safety, the research is focused on the safety aspects involving radiation in medical imaging. The IAEA's Activities on Radiation Protection in Interventional Cardiology by Madan (2007) is one of the examples for safety-radiation safety research title. For clinical-modality, the research area is focused on the modalities in medical imaging that is used in clinical area. An example of the research is Carcinoma of stomach detected by routine transabdominal ultrasound (Wong et al*.,* 2010)

For non-Malaysian authors, the significant themes and sub-themes were clinical-diagnosis, clinical-health technology assessment and modality-technical. For collaboration works, the most significant theme and sub-theme was clinical-health technology assessment. Thus, it can be concluded that most authors focus on clinical themes in their researches.

The issue associated with safety is among the important themes being discussed by researchers with 33 articles out of 268 articles (12.3%). Radiation safety was the most frequent topic to be discussed by authors. According to Ploussi and Efstathopoulos (2016), the higher usage of ionization radiation for diagnostic and therapeutic purposes has raised significant safety and health concerns for patients and medical imaging practitioners. As the number of medical imaging procedures performed continues to rise each year, there should be a precaution to keep radiation dose as low as reasonably achievable (Marshall & Keene, 2006). Thus, the safety aspect in medical imaging field is considered as a significant area as it can contribute to other related risks of radiation such as cancer. The needs of safety aspect to be constantly observed cannot be denied. Hence, aspects that involves radiation safety is an important aspect to be continuously researched.

The results suggest that the themes that have the lowest frequencies from both Malaysian and non-Malaysian authors were economy and quality. In medical imaging, cost considerations, quality and patient-centered researches would also be highly relevant. It was also observed that no research based on patient-centered has been done such as patient safety, patient preparation, patient education and patient care. In order to deliver the optimal service, attention should also be given to these aspects.

CONCLUSION

This study established the trend of research in medical imaging as featured in the MyCite database. It can be deduced that the contributions by medical imaging practitioners, both academicians and service providers as indexed in MyCite, is low. Practitioners should also take note of the presence of journals indexed in MyCite that are also indexed in Scopus and WoS. This is to reduce the conflict between publishing in high impact foreign journals and those local journals indexed in the two renowned databases. Most of the authors are of university-based and hospital university-based background while the involvement of authors from the hospital background is low. The thematic analysis showed that clinical studies formed the main preference among researchers. This study suggests that researchers should also pay attention to non-clinical themes such as economic, safety, and education. Studies need to be conducted to determine the status of involvement in research by graduates in medical imaging who are now serving the hospitals. Furthermore, the involvement in research among medical imaging practitioners and academicians should be directed to preferences of research that can address the local needs. This particular work can be a guideline to future researchers in order to decide on their future research undertakings.

ACKNOWLEDGEMENT

A bundle of thanks to those who contribute to this research paper directly and indirectly. It is hoped that this research paper may benefit researchers and students in research preferences.

REFERENCES

Abdul, K. A. H., Win, M., & Izaini, G. A. R. (2010). Computed tomography perfusion imaging on traumatic cerebral contusion: a preliminary report. *The Malaysian Journal of Medical Sciences, 17*(4), 51-56.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77-101.

Braun, V., & Clarke, V. (2014). What can "thematic analysis" offer health and wellbeing researchers? *International Journal of Qualitative Studies on Health and Wellbeing, 9*(1), 1-2.

Bruns, C. (2010). *Empirical research: how to recognize and locate*. Paulina June and George Pollak Library, California State University. Retrieved March 19, 2017, from http://users.library.fullerton.edu/cbruns/empirical_research.htm

Clarke, A., Gatineau, M., Grimaud, O., Royer-Devaux, S., Wyn-Roberts, N., Bis, I. L., & Lewison, G. (2007). A bibliometric overview of public health research in Europe. *European Journal of Public Health, 17*(1), 43-52.

Economic Planning Unit. (2010). *Tenth Malaysia Plan 2011-2015*. The Economic Planning Unit, Prime Minister's Department Putrajaya, Malaysia. Retrieved January 10, 2017, from https://www.pmo.gov.my/dokumenattached/RMK/RMK10_Eds.pdf

Ethnographic Medical Research Group (EMeRG). (2016). *How would the future of Malaysian private hospitals shape?* Ethnographic Medical Research Group (EMeRG). Retrieved June 20, 2017, from https://www.slideshare.net/EMeRG-research/review-of-private-hospitals-in-malaysia

Fathinul, F. A. S., & Abdul, J. (2012). Combined ultrasound and IVU for the management of childhood urolithiasis: a case report. *Pertanika Journal of Science and Technology, 20*(2), 327-330.

Hasnah, H., Amin, I., & Suzana, S. (2012). Bone health status and lipid profile among post-menopausal Malay women in Cheras, Kuala Lumpur. *Malaysian Journal of Nutrition, 18*(2), 161-171.

Kai, X., Sooi, H. O., & Aboul, E. H. (2007). Brain magnetic resonance image lateral ventricles deformation analysis and tumor prediction. *Malaysian Journal of Computer Science, 20*(2), 115-132.

Kamal, I., Kumari, C. K., Vijayalakshimi, K., & Tamanang, S. (2015). Determination of tube output (kVp) and exposure mode for breast phantom of various thickness/glandularity for digital mammography. *The Malaysian Journal of Medical Sciences, 22*(1), 40-49.

Madan, R. (2007). The IAEA's activities on radiation protection in interventional cardiology. *Biomedical Imaging and Intervention Journal, 3*(2), 1-5.

Marshall, G., & Keene, S. (2006). Radiation safety in the modern radiology department: a growing concern. *The Internet Journal of Radiology, 5*(2), 1-6.

Mohamad, N. S., Nazimah, I., & Rushdan, M. N. (2012). The role of computed tomography (CT) scan in assessment of the parametrial involvement in early stage cervical carcinoma. *Journal of Surgical Academia, 2*(1), 14-17.

MyCite. (2017). *Ministry of Higher Education*. Retrieved November 15, 2016, from http://www.mycite.my/

Nor, A. M., & Zainul, A. R. (2015). Prosthetic ear reconstruction applying computer tomographic (CT) data and additive manufacturing technologies. *Jurnal Teknologi, 76*(7), 63-68.

Noriah, J., Humairah, S. C., Siti, K. C. M., & Ellyda, M. N. (2013). A survey on physical factors and compressed breast thickness in voluntary mammography screening using FFDM system in Malaysia: preliminary results. *Jurnal Sains Nuklear Malaysia, 25*(1), 30-35.

Philip, M. (2009). *Guideline for writing a review article*. Retrieved June 20, 2017, from http://www.plantscience.ethz.ch/education/Masters/courses/Scientific_Writing

Ploussin, A., & Efstathopoulos, E. P. (2016). Importance of establishing radiation protection culture in radiology department. *World Journal of Radiology, 8*(2), 142-147.

Ponto, J. A. (2015). Understanding and evaluating survey research. *Journal of the Advanced Practitioner in Oncology, 6*(2), 168-171.

Reeves, P. J. (2008). Research in medical imaging and the role of the consultant radiographer: a discussion. *Radiography, 14*, e61-e64.

Rose, S., Spinks, N., & Canhoto, A. I. (2015). *Management research: Applying the Principles*. New York, NY: Routledge.

Snaith, B. (2012). Collaboration in radiography: A bibliometric analysis. *Radiography, 18*(4), 270-274.

Suryani, I., Yaacob, A., Hashima, N., Rashid, S. A., & Desa, H. (2013). Research publication output by academicians in public and private universities in Malaysia. *International Journal of Higher Education, 2*(1), 84-90.

Wong, M. F. E., Shum, S. F. J., & Chau, W. K. (2010). Carcinoma of stomach detected by routine transabdominal ultrasound. *Biomedical Imaging and Intervention Journal, 6*(4), 1-3.

Yartsev, S., Kron, T., & Van, D. J. (2007). Tomotherapy as a tool in image-guided radiation therapy (IGRT): theoretical and technological aspects. *Biomedical Imaging and Intervention Journal, 3*(1), 1-10.

Zainab, A. N., Sanni S. A., Edzan, N. N., & Koh, A. P. (2012). Auditing scholarly journals published in Malaysia and assessing their visibility. *Malaysian Journal of Library and Information Science*, 17(1), 65-92.

Pertanika J. Sci. & Technol. 27 (4): 1485 - 1508 (2019)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Review article

Psychosocial Predictors of Adolescent Aggression

Fatimah Ahmad Fauzi, Nor Afiah Mohd Zulkefli*, Halimatus Sakdiah Minhat and Norliza Ahmad

Department of Community Health, Faculty of Medicine and Health Sciences, Universiti Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia

ABSTRACT

Aggression is one of the most common Paediatric Psychiatric emergency problems presented in healthcare services with major public health impacts. The aim of this review is to determine the psychosocial predictors of adolescent aggression. Systematic review on observational study designs were conducted. Recent five years English published journal articles (2014-2018) were searched in three databases (Science Direct, PubMed and EBSCO) from April 2018 until May 2018. The final 15 articles (seven prospective cohort, eight cross-sectional studies) were included in the systematic review. Psychological predictors were mainly on the personality traits and emotional problems experienced by the adolescents including callous-unemotional traits, hostility, anger, and low empathy. Self-control was found to be protective against adolescent aggression. Social predictors that contribute to the aggression among adolescents according to the highest reported number of articles were peer influence, followed by school climate, substance use, neighbourhood influence, family, and parent factors. In conclusion, psychosocial predictors of adolescent aggression include both Psychological Factors; mainly personality traits and emotional problems, and Social Factors; mainly peer influence and substance use. This warrants for more holistic approach in dealing with aggressive adolescents that can be applied in more targeted and focused intervention strategy deliveries.

Keywords: Adolescent, aggression, predictors, psychosocial

ARTICLE INFO

Article history:
Received: 05 July 2018
Accepted: 13 May 2019
Published: 21 October 2019

E-mail addresses:
fatimahfauzi@yahoo.com (Fatimah Ahmad Fauzi)
norafiah@upm.edu.my (Norafiah Mohd Zulkefli)
halimatus@upm.edu.my (Halimatus Sakdiah Minhat)
lizaahmad@upm.edu.my (Norliza Ahmad)
* Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

INTRODUCTION

The term “Aggression” is always used interchangeably with aggressive behaviour, which can be defined as any action with the purpose of resulting harm, injury, or pain against others (Zirpoli, 2008). It is a broad construct that comprises physical such as at-risk behaviours, delinquency, violence, and non-physical forms such as bullying, slandering, and suicide behaviours (Liu et al., 2013). There are many classifications of aggression (Liu, 2004) based on its forms (overt and relational), motives (reactive and proactive), and functions (hostile and instrumental). Overt aggression involves obvious and outward confrontational acts of aggression, which differ from the covert or relational aggression that is more hidden and manipulative (Conner & Barkley, 2004). Reactive and hostile aggression reflects overheated and uncontrolled reactions in response to potential harm caused by physical or verbal aggression initiated by others as to protect oneself, whereas proactive and instrumental aggression reflect controlled, purposeful, and cold-blooded actions in achieving one's goal without harm consideration for others (Liu, 2004; Card & Little, 2006). These typologies are subjected to overlap, comprising of physical and verbal aggressive acts. Though this review is on adolescent aggression, the literature review also covers related disruptive and criminal behaviour, delinquent, bullying as well as violence.

Adolescence is being perceived as a period of "storm and stress" (Hamid et al., 2015), which predisposes them into more serious aggression and violence in the early period including gang fights and use of knives (Liu, 2004). It becomes a public health concern as it contributes to the major cause of mortality in this age group by risk-taking behaviours that can result in unintentional injuries (Rashtriya & Swasthya, 2014; World Health Organization, 2017). In the year 2000, the violent death rate in low- to middle-income countries was 32.1 per 100 000 population, which was more than twice as compared to the rate in high-income countries of 14.4 per 100,000 population (Krug et al., 2002). This disproportionate statistic persisted as homicide rate worldwide report in 2012 was more pronounced in low- and upper middle-income countries than in lower middle- and high-income countries (Butchart & Mikton, 2014).

Following the 2002 World report on violence and health, many programmes, policies and legislative measures have been implemented across the worlds to prevent violence-related aggression in the community (Butchart & Mikton, 2014). The strategies proposed for the children and adolescents included developing life skills, safe, stable and nurturing relationships with their parents and caregivers, and reducing the common factors of violence act such as alcohol abuse and access to weapon. However, aggression was still prevalent among youth, with the highest homicide rates in the world in 2012 represented by adolescents' aged 15-29 years old; 10.9 per 100,000 population. As in Malaysia, aggressive-related behaviours such as bullying and physical fight were reported to occur amongst 28% of the adolescents (Hussin et al., 2014) and the juvenile crime cases reported among school students included homicide, rape, robbery, wilderness, and inflicting injury to others (Ibu Pejabat Polis Kontijen, 2018).

Different socio-cultural influence of aggression manifestation warrants for different strategical approaches (Kim et al., 2010). Hence, interventions succeeded in developed countries may not be practical in other developing countries. Even though many interventions preventing adolescent aggression conducted in Malaysia, which include health education, counselling, and rehabilitation programmes (World Health Organization, 2006), their effectiveness and evidence-based driven are still not well established and informed.

Aggression is one of the most common paediatric psychiatric emergency problems presented in healthcare services (Carubia et al., 2016). It can be associated with many psychiatric disorders such as anxiety (Liu et al., 2013) and depression that was reported to occur amongst 10-20% of Malaysian adolescents (Srinath et al., 2010). Increased aggression in adolescence may also indicated atypical development and higher vulnerability for future negative mental health outcomes (Yang et al., 2016).

Psychosocial relates to "the interrelation of social factors and individual thought and behaviour" (Psychosocial, 2019, para 2). It has important implications for health researchers and social epidemiologists as psychosocial determinants of health can explain the mediating effects and contextual factors of social structures on individual's health (Martikainen et al., 2002). Many studies have investigated psychological factors of adolescent aggression such as self-esteem (Arokiaraj et al., 2011), emotional intelligence (Masoumeh, 2014), antisocial personality (Duru, 2015), and social factors including family environment (Seong, 2008; Azmawati et al., 2015) and functioning (Zainah et al., 2011), peer and teacher attachment (Duru, 2015). However, these underlying factors were looked into the psychosocial contribution separately, which is important to be collectively integrated in public health action (Macleod & Smith, 2003). Psychosocial treatments have shown promising effect on aggressive children and adolescents with conduct disorder (Kazdin, 1997). This warrants for a systematic review to determine the common psychosocial predictors of aggression among adolescents, joint in one review, and compare psychological and social predictors that were commonly reported from the recent research evidences. In that way, evidence-based practice in developing more targeted and effective preventive interventions can be produced to curb this global situation of aggression among adolescents.

MATERIALS AND METHODS

A systematic review on observational study designs (cohort, case-control, and cross sectional studies) were conducted using the Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols (PRISMA-P) 2015 statement (ref:http://www.prisma-statement.org) and registered with PROSPERO (registration no: CRD42018093821). Recent five years published journal articles (2014-2018) were searched in three databases, incorporating medical and social science literature (Science Direct, PubMed, and EBSCO) from April 2018 until May 2018. Only English language articles were searched by the researchers with consultation from the university health sciences librarian, using search techniques of Boolean operator, truncation & phrase searching of specific words and MeSH terms: (psychological OR social) AND (predictors OR determinants OR factors) AND (adolescent* OR teenager* OR youth) AND (aggression OR agonistic behaviour OR bullying). As the systematic review focusing on 'aggression' term, only one source of MeSH terms (PubMed) was used, excluding other related MeSH terms such as juvenile delinquency and violence.

Considering general classification of aggression and its detrimental health impact on adolescents, the focus of the review was on studies that investigated physical and/or verbal aggression. Self-reported validated questionnaires that measure adolescent aggression by continuous means or categorical were included in the review, as there is no gold standard measurements on adolescent aggression yet. Studies in both developed and developing countries, which addressed bias and efforts taken to reduce them were also included in the review. Other systematic reviews or review articles and non-empirical works such as case studies or commentaries were excluded from the review. For each selected study, the participants need to be randomly selected as to represent the adolescents age (10-19 years old according to WHO standard) or secondary/middle school students. As the review focused on psychosocial predictors of adolescent aggression, study participants with underlying organic brain injury, psychiatric disorder, or neurological disorder that can manifest aggression as one of the symptoms will be excluded. For cohort studies, at least two follow-up points were required for the results to be included in the review. As for case-control studies, the cases and controls done should derive from similar population background and characteristic. If they were matching, the cases and controls should not be matched more than 1:3 ratio. The statistical analyses for each study designs should apply multivariate analysis in eliciting the significant predictors and presented in Adjusted Odds Ratio (AOR), Relative Risk (RR), or Beta coefficient with 95% Confidence Interval not crossing at one or p value less than .05. Cross-sectional studies that met the inclusion and exclusion criteria were only included in the systematic review.

The process of study selection was conducted in two stages of (1) primary screening on the title/abstract of the articles, and (2) secondary screening on the full text articles when inclusion or exclusion criteria were not clear from the title/abstract. Article citations were organized, downloaded and reviewed in Mendeley desktop. After duplication of articles were recognized by comparing the author names and study title, selected full texts were retrieved and systematically assessed to be included in the review. The eligibility criteria screening for study inclusions was conducted by one reviewer, followed by assessment of second set of reviewers (N.A.M.Z., H.S.M, and N.A) on the retrieved study articles for any doubt on article's eligibility. Disagreements were met by mutual consensus.

At present, there is still lack of specific tool to assess the quality of aggression studies. Hence, the NIH (National Institutes of Health) Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies was used to assess the quality of study design, data collection, and data analysis. Various results obtained were classified into psychological and social predictors with different study designs in a logical framework of each study's overview of author/year, participants/inclusion criteria, outcome measurements, and results of a study. After summarizing all the selected research evidence, which included tabulation of study characteristics and quality, they were interpreted and recommended for improvements based on their limitations in the discussion. Due to the high heterogeneity of included observational study designs and their outcome measures, meta-analyses are unable to be computed for this review.

RESULTS

The electronic search through three databases produced initial 670 references (PubMed= 545; EBSCO= 107; Science Direct= 18). The search process required assessment of the title, proceed by the abstract. Eight articles were found to be duplicated across all the searches, which resulted in 662 articles to be screened after duplications being removed. Most of the articles; 638 out of 662, were excluded from the primary screening on title and abstract as they didn't fulfil the study selection criteria such as more than five-years recent articles and unpublished observational studies. After the remaining articles were screened based on inclusion criteria, only 24 full articles were retrieved for secondary screening and assessments of the articles' content, which covered the introduction, methodology, result, discussion, and conclusion. A total of nine articles were excluded due to several reasons: absence of regression analysis findings, only one-point follow up in cohort study, youth study participants, and aggression as the independent rather than dependent variable. This has resulted in the final 15 articles (seven were prospective cohort studies, eight were cross-sectional studies) that were included in the systematic review. There was a lack of recent case-control studies that met the inclusion criteria to be included in this systematic review. The flow diagram of study selection is described in Figure 1, guided by the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) guidelines 2009 (Moher et al., 2009).

Study quality assessment using the NIH Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies showed fair to good quality of the selected studies (Table 1). All studies described the study objectives clearly with participation rate of eligible person of more than 50% except for several studies among urban African American (Finigan-Carr et al., 2015), adolescents with severe conduct problems in Amsterdam (Jambroes et al., 2018), and Positive Action's participants in North Carolina (Stalker et al., 2018). The independent variables were clearly defined, valid, reliable, implemented

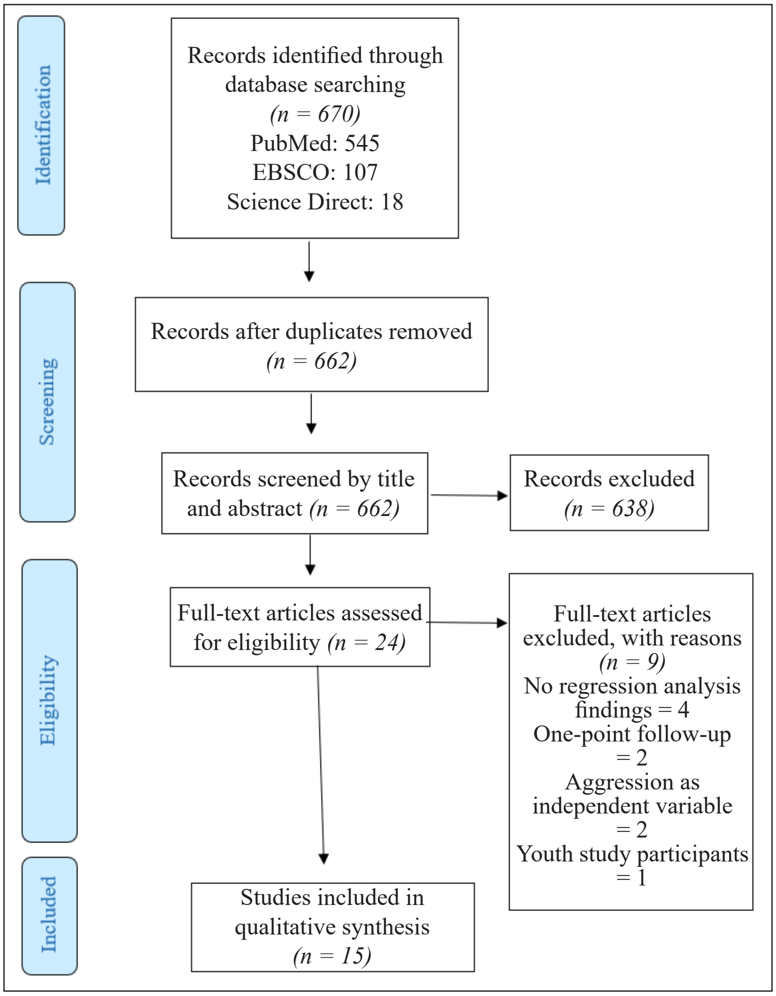


Figure 1. PRISMA flowchart for systematic search on psychosocial predictors of adolescent aggression

constantly across all study participants, and assessed more than once over time. Same goes for the dependent variables which encompassed physical and/or verbal aggression, have been validated and presented good reliability test (Cronbach's α >0.70) in respective population as shown in Tables 2 and 3. Confounding variables were also being measured and adjusted by statistical measures in determining the relationship between independent and dependent variables.

The findings of each article were discussed in Table 2 (cross-sectional study) and Table 3 (cohort study). There were eight cross sectional studies with participants ranging from 156 (Barry et al., 2018) to 4,674 adolescents (Kivimaki et al., 2014), mainly in developed countries except for single study in developing (China) and undeveloped country (Gaza). Almost all studies examined adolescents in public schools except for several studies which focused on adolescents with conduct problems like school dropped-out (Barry et al., 2018;

Table 1
Quality assessment tool for studies of adolescent aggression predictors

Criteria	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. Research question/ objective clearly stated	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
2. Study population clearly specified & defined	√	X	X	√	√	√	√	√	X	X	X	√	X	√	√
3. Participation rate at least 50%	X	√	√	X	√	NR	√	√	√	√	√	√	√	X	√
4. Subjects recruited from same population with uniformly applied eligibility criteria	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
5. Sample size justification, power description, or variance and effect estimates	X	X	X	X	X	X	X	X	X	X	√	X	X	√	X
6. Exposure(s) of interest measured prior to the outcome(s)	X	X	X	X	X	X	X	X	√	√	√	√	√	√	√
7. Timeframe sufficient to see an association between exposure and outcome	X	X	X	X	X	X	X	X	√	√	√	√	√	√	√
8. Examine different levels of the exposure as related to the outcome	NA	NA	√	NA	√	NA	√	√	√	√	√	√	√	√	√
9. Independent variables clearly defined, valid, reliable, and implemented consistently	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
10. Exposure(s) assessed >1 over time	NA	NA	NA	NA	NA	NA	NA	NA	√	√	√	√	√	√	√
11. Dependent variables clearly defined, valid, reliable, and implemented consistently	NR	√	√	√	√	√	√	√	√	√	√	√	√	√	√
12. Outcome assessors blinded to the exposure status of participants	√	NA	NA	NR	√	NR	√	NR	X	X	√	X	√	√	√
13. Loss to follow-up after baseline 20% or less	NA	NA	NA	NA	NA	NA	NA	NA	√	√	√	√	√	X	X
14. Key potential confounding variables measured and adjusted statistically	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√

Note. NR = not recorded; NA= not applicable

Table 2
Psychosocial predictors of adolescent aggression in cross-sectional study articles

Author/Year	Participants/Inclusion criteria	Outcome Measurement	Results
1. Title: Using the Theory of Planned Behavior to Predict Aggression and Weapons Carrying in Urban African American Early Adolescent Youth			
Finigan-Carr, Cheng, Gielen, Haynie, & Simons-Morton (2015) https://doi.org/10.1177/1090198114548479	452 of 6th graders in three urban middle schools on probation for classification as "persistently dangerous" under the State of Maryland's No Child Left Behind Act policy. The sample was mainly African American, with 12.6% Hispanic. There were almost equal numbers of males and females with a median age of 12 years old.	• Frequency of aggressive behaviours were measured by five items developed by Bosworth and Espelage (1995) to gather information about physical and nonphysical aggressive behaviours	• Those with medium and high levels of **problem friend influence** were found to be almost three (AOR= 2.61; 95%CI= 1.68-4.07) and five times (AOR= 5.05; 95%CI= 3.03-8.39) more likely to manifest aggressive behaviors, respectively. • Adolescent reports of **parental disapproval of fighting** (AOR= 0.39, 95%CI= 0.21-0.76) and **increased self-control** (AOR= 0.59; 95%CI= 0.39-0.88) were seen as protective factors.
2. Title: The Indirect Effects of Adolescent Psychopathic Traits on Aggression Through Social-Cognitive Factors			
Lui, Berry, & Schoessler (2017) https://doi.org/10.1007/s10826-017-0667-y	209 adolescents (178 males, 31 females) aged 16 to 19 years (M=16.83, SD=.80), from a voluntary military-style residential program in south-eastern USA for school dropped-out adolescents. Ethnicities: 56% White, 34.4% Black, 1% other ethnicities (8.6%)	• Peer Conflict Scale. Factor analysis confirmed existence of reactive and proactive aggression dimensions, Cronbach's alpha was 0.83 for reactive aggression and .88 for proactive aggression (Marsee et al., 2011).	• The model with **CU traits** (callous, uncaring, unemotional, narcissism) predicting reactive (R^2= 0.24; b= .06; 95%CI= -0.02 - 0.15) and proactive aggression (R^2= 0.14; b=.14; 95%CI= -0.01 - 0.16) had a significant total effect.
3. Title: Global and contingent self-esteem as moderators in the relations between adolescent narcissism, callous unemotional traits, and aggression			
Barry, McDougall, Anderson, & Bindon (2018) https://doi.org/10.1016/j.paid.2017.10.036	156 adolescents (126 males, 29 females, 1 unreported; 81% males, 19% females) aged 16–19 (M = 16.81, SD = 0.77), enrolled in a voluntary military-style program for at-risk youth who have dropped out of school. The racial/ethnic composition was 52.6% White, 29.5% Black, 0.6% Hispanic, 0.6% Other, and 16.7% unreported.	• Peer Conflict Scale (PCS; Marsee et al., 2011) assess reactive and proactive aggression. Good reliability with Cronbach's α value of 0.95 for proactive aggression and 0.90 for reactive aggression.	• Significant main effect for **narcissism** in predicting self-reported proactive aggression, b = 0.16, se = 0.05, $p < 0.001$, and reactive aggression, b = 0.14, se =0.05, p= 0.003. • significant main effects for **CU traits** in the prediction of self-reported proactive aggression, b = 0.42, se = 0.09, $p < 0.001$, and self-reported reactive aggression, b = 0.35, se = 0.09, $p < 0.001$.

Table 2 *(continue)*

Author/Year	Participants/Inclusion criteria	Outcome Measurement	Results
4. Title: Dimensions of psychopathy in relation to proactive and reactive aggression: Does intelligence matter?			
Jambroes et al. (2018) https://doi.org/10.1016/j.paid.2018.03.001	159 adolescents between the ages of 12 and 18years old (mean age 15.1± 1.3 years), who attended closed treatment (TACt, Training Aggression Control) for severe conduct problems in Amsterdam. Almost half (49%) was male, and 37% was of non-Western origin (predominantly Moroccan or Surinamese).	Dutch version of the Reactive and Proactive aggression Questionnaire (RPQ). Good reliability with Cronbach's α of 0.90 for proactive and reactive aggression scale (Cima, Raine, Meesters, & Popma, 2013)	• **Callous-unemotional** dimension of psychopathy predicted reactive (β= 0.26; 95%CI= 0.02-0.36) and proactive aggression (β= 0.25; 95%CI= 0.02-0.36) • **Impulsive-irresponsible** dimension of psychopathy predicted reactive (β= .36; 95%CI= .08-.35) and proactive aggression (β= .25; 95%CI= 0.02-0.29)
5. Title: Predicting Aggression in Adolescence: The Interrelation between (a lack of) Empathy and Social Goals			
Van Hezebroek, Olthof, & Goossens (2016) https://doi.org/10.1002/ab.21675	550 participants (49.5% boys), between age 11 and 14 (mean age 12.97 years, SD= 5.69 months) from five schools (25 classes) in Netherland.	Instrument for Reactive and Proactive Aggression (IRPA) (Polman, Orobio De Castro, Thomaes, & Van Aken, 2009) measured physical, verbal, hidden and relational aggressive behaviour. Good reliability for reactive (Cronbach's α = 0.82) and proactive (Cronbach's α = 0.84) aggression.	• **lack of empathic** significantly predicted both proactive (b= 0.005; p<0.01) and reactive (b= 0.007, p= 0.02) aggression • **agentic goals** significantly predicted proactive aggression (b= 0.026; p<0.01) • model R^2 = 0.04
6. Title: School climate and adolescent aggression: A moderated mediation model involving deviant peer affiliation and sensation seeking			
Wang, Yu, Zhang, Chen, Zhu, & Liu (2017) https://doi.org/10.1016/j.paid.2017.08.004	1401 early adolescents (50.2% male) ranging from 11-14 years old (mean:12.46 ± 0.61 years) from 4 junior middle schools in Guangdong. 21.3% came from rural areas, 24.2% from county seats, 27.8% from small-medium cities, and 26.6% from metropolitan areas. Most students' family monthly income (71.7%) between ¥ 1000 - ¥ 4000 with less than junior college educated parents.	Chinese version of the Buss-Warren aggression questionnaire (BWAQ; Maxwell, 2008) measures physical, verbal, and indirect aggression. Cronbach's α = 0.86.	• High quality **school climate** negatively predicted adolescent aggression (*b*= -0.21; p<0.01) • **Deviant peer affiliation** predicted adolescent aggression (*b*= 0.18; p<0.01)

Table 2 *(continue)*

Author/Year	Participants/Inclusion criteria	Outcome Measurement	Results
7. Title: Positivity Ratio Links Self-control Skills to Physical Aggression and Happiness in Young Palestinians Living in Gaza			
Rosenbaum, Ronen, Abuelaish, Orkibi, & Hamama (2018) https://doi.org/10.1177/1103308817743371	744 Gazan Palestinian adolescents, 48% girls (n = 358). Participants were born between 1988 and 1995, with mean age of 15.67 years (SD = 1.62), studying in Grades 8–12 (part of compulsory basic and secondary education in Gaza). As for religion, 96% were Muslim, and 4% were Christian, Druze or Bedouin.	Buss and Perry's (1992) Aggression Questionnaire, which includes four subscales: hostile thoughts, anger, verbal aggression and physical aggression. The Arabic version scale's internal consistency coefficient value ranged 0.63-0.77 (acceptable).	• **Self-control skills** negatively predicted adolescent aggression (β= -0.07; p<0.05) • **Hostility** (β= .14; p<0.001) and **anger** (β= 0.14; p<0.001) directly predicted adolescent aggression
8. Title: Alcohol use among adolescents, aggressive behaviour, and internalizing problems			
Author/Year	**Participants/Inclusion criteria**	**Outcome Measurement**	**Results**
Kivimäki et al. (2014) https://doi.org/10.1016/j.adolescence.2014.06.011	4074 pupils of 13-18 years of age from all local schools (excluding schools for pupils with impaired cognitive skills) in Kuopio, Eastern Finland. 53% of them were female and the types of school were comprehensive school for 1840, upper secondary school for 1474 and vocational school for 900 participants.	ASEBA- Youth Self report (Achenbach & Rescorla, 2001), which examined multiple aspects of adolescent behaviour and well-being including aggressive behaviour. Cronbach's α value ranged from 0.7 - 0.83.	• **Alcohol use** predicted aggressive behaviour in both male (OR= 1.09; 95%CI: 1.05-1.13) and female (OR= 1.11; 95%CI= 1.06-1.16) • **Level of alcohol consumption** predicted aggressive behaviour in both male (IRR= 1.03; 95%CI= 1.02-1.04) and female (IRR= 1.02; 95%CI= 1.01-1.02)

Note.
TPB = theory of planned behaviour
CU = callous-unemotional
ASEBA= Achenbach System of Empirically Based Assessment

Table 3
Psychosocial predictors of adolescent aggression in cohort study articles

Author/Year	Participants/Inclusion criteria	Outcome Measurement	Results
9. Title: Aggressive Delinquency Among North American Indigenous Adolescents: Trajectories and Predictors			
Sittner & Hautala (2016) https://doi.org/10.1002/ab.21622	659 adolescents from waves 1, 2, 3, 5, and 7 of a longitudinal study on three reservations in the northern Midwest and four Canadian reserves.	• Aggressive delinquency measured by nine items, adapted from conduct disorder module of the Diagnostic Interview Schedule for Child IV. Fair to good reliability with Cronbach's α = 0.68 - 0.73 (Shaffer, Fisher, Lucas, Dulcan, & Schwab-Stone, 2000).	• **Increased parental rejection** associated with higher risk of being chronic desistor (RR= 1.36, p<0.05) • **Positive school adjustment** associated with aggression trajectory profile (RR= 0.65 in high desistor and RR=0.67 in chronic desistor, p<0.05) • **Delinquent peer associations** increased the risk of being in high desistor (RR= 2.05, p<0.01) and chronic groups (RR= 2.94, p<0.01) • **Ever trying substances** increased risk of being in moderate desistor (RR= 2.45), high desistor (RR=3.45), and chronic desistor (RR=4.78) • **Early dating** increased the relative risk of being in the high desistor (RR= 2.76; p<0.05) and chronic groups (RR= 2.65; p<0.05) than in the non-offender group
10. Title: Multi-level risk factors and developmental assets associated with aggressive behaviour in disadvantaged adolescents**.**			
Smokowski, Guo, Cotter, Evans, & Rose (2016) https://doi.org/10.1002/ab.21612	Middle school students came from the RAP study and the data were collected in spring of 2011, spring of 2012, and spring of 2013. 4,065 observations at baseline, 4,251 observations of Wave 2 or 12 months after the baseline, and 4,256 observations at Wave 3 or 24 months after the baseline. They were racially/ ethnically diverse: 27% White, 23% African-American, 28% American-Indian, 14% mixed race/other, and 8% Latino. About half of the sample (52%) was female, 85% of resided in a two-parent family, and 86% received free or reduced-price lunch.	Modified subscales from the YSR. Good reliability with Cronbach's alpha value 0.86 -0.70 for each year of follow-up (Achenbach & Rescorla, 2001)	• **Two-parent family structure** predicted aggression (ExpB= 0.979, p<.001) • Negative proximal processes of **parent-adolescent conflict** (ExpB= 1.013, p<0.001), **friend rejection** (ExpB= 1.023, p<.001), **peer pressure** (ExpB= 1.018, p<0.01), **delinquent friends** (ExpB= 1.124, p<0.001), and **school hassles** (ExpB= 1.063, p<0.001) were significant predictors of aggression. • Positive proximal processes of **ethnic identity** (ExpB= 0.991, p<0.01), **religious orientation** (ExpB= 0.98, p<0.001),and **school satisfaction** (ExpB= 0.948, p<0.05) buffers against aggression.

Table 3 *(continue)*

Author/Year	Participants/Inclusion criteria	Outcome Measurement	Results
11. Title: Social Information Processing in Child-to-Parent Aggression: Bidirectional Associations in a 1-Year Prospective Study			
Calvete, Gámez-Guadix, García-Salvador(2015) https://doi.org/10.1007/s10826-014-0023-4	Participants were students from 67 classrooms located within 21 secondary schools in Bizkaia, Basque Country, Spain, including public and private centres. Participants were 1,272 adolescents (653 girls, 619 boys; Mean age = 14.74, SD = 1.21), who completed the measures at two- time measurements within a 1-year interval. Majority of participants were Spanish (91.5%), and the remaining participants were South American (4.9%), Eastern European (0.6%), African (0.5%), and Asian (0.2%).	• Child-to-Parent Aggression Questionnaire (CPAQ; Calvete et al., 2013). Excellent psychometric properties in a Spanish adolescents, with exploratory factor support for its factor structure and good reliability (Cronbch's α = 0.75-0.85)	• Aggressive **response access** (β= 0.18, $p < 0.001$) and **anger** (β= 0.08, $p < 0.01$) predicted T2 psychological aggression • Anger (β= 0.07, $p < 0.05$), **low empathy** (β= -0.08, $p < 0.05$), and **hostile attribution** (β= 0.08, $p<0.05$) predicted T2 physical aggression. • Model explained 12% variance of psychological and physical aggression.
12. Title: Substance Use and Physical Dating Violence: The Role of Contextual Moderators			
Reyes, Foshee, Tharp, Ennett, & Bauer (2015) https://doi.org/10.1016/j.amepre.2015.05.018	2,299 public school students in two counties of North Carolina contributed at least two waves of data, with 78% participating in three or more waves (n=1,920). The participants' demographic characteristics were 47% black, 48% were male, and 40% with either parent's highest education was high school or less.	Physical violence dating perpetration was measured by six items listing physically violent behaviour acts. Good reliability with Cronbach's α of 0.93 (Reyes et al., 2015)	• Physical dating violence perpetration increased at time points when **heavy alcohol** (β coefficient= 0.17, $p<0.001$) and **hard drug use** were elevated (β coefficient= 0.05, $p=0.01$)
13. Title: The Role of Bicultural Stress and Perceived Context of Reception in the Expression of Aggression and Rule Breaking Behaviours Among Recent-Immigrant Hispanic Youth			
Forster, Grigsby, Soto, Schwartz, & Unger (2015) https://doi:10.1177/0886260514549052	302 adolescents recently arrived (5 years or less) Hispanic immigrants in Miami and Los Angeles and participated in COPAL study. Participants from Los Angeles were mainly Mexican (70%), and those from Miami were mainly Cuban (61%). Adolescents were entering or currently enrolled in 9th grade at a public school in 10 schools in Miami-Dade or 13 schools in Los Angeles Counties. Only schools that were at least 75% Hispanic were selected for the study.	Aggressive Behaviour subscale from the Youth Self-Report. Good reliability with Cronbach's α of 0.91 (Achenbach & Rescorla, 2001)	• **Bicultural stress** (β = 0.318, $p < 0.001$), **negative context of reception** (β = .184, p = .021), aggression at baseline (model R^2 = 0.299; β = 0.220, p = 0.027) • **Delinquent peers** (β = 0.1024, p = 0.047) had independent associations with changes in aggressive behaviour during the first year of high school (model R^2 = 0.37, $p < 0.001$).

Table 3 *(continue0*

Author/Year	Participants/Inclusion criteria	Outcome Measurement	Results
14. Title: The impact of the positive action program on substance use, aggression, and psychological functioning: Is school climate a mechanism of change?			
Author/Year	**Participants/Inclusion criteria**	**Outcome Measurement**	**Results**
Stalker, Wu, Evans, & Smokowski (2018) https://doi.org/10.1016/j.hildyouth.2017.11.020	8333 participants participated in Rural Adaptation Project (RAP) in North Carolina; 50.63% were female. Racial/ethnic compositions: 28.27% Caucasian, 25.50% African American, 25.35% American Indian, 12.26% Mixed Race/Other, and 8.62% Hispanic/Latino. Majority (78.99%) received free/reduced price lunch and 70.32% lived with two parent families.	Youth Self-Report. Good reliability with Cronbach's α of 0.91 (YSR; Achenbach & Rescorla, 2001)	• **School hassles** (as proxy to school climate) predicted adolescent aggression (b= 0.347; p<0.001)
15. Title: Gender moderates the association between psychopathic traits and aggressive behavior in adolescents			
Orue, Calvete, & Gamez-Guadix (2016) https://doi.org/10.1016/j.paid.2016.01.043	765 (464 girls and 301 boys) high school students from 13 randomly selected schools in Bizkaia (Spain) completed the measures at both Time 1 (T1) and Time 2 (T2) one year later. The participants' age between 14 -18 years old (Mean = 15.43 ±1.09 years). Most of the participants were Spanish (90.1%) or South American (7.9%). Remaining 2% were from various countries. The socio-economic levels were low (12.1%), low-medium (17.7%), medium (32.7%), high-medium (29.9%), and high (7.6%).	Self-report questionnaire developed by Little, Henrich, Jones, and Hawley (2003), which differentiates the forms and functions of the aggressions as proactive overt, proactive relational, reactive overt and reactive relational aggression. Fair to good reliability test (Cronbach's α = 0.67 - 0.76)	• T1 **CU traits** predicted T2 proactive overt (b= 0.07; p<0.05) and proactive relational (b= 0.10; p<0.05) aggressive behavior. • T1 **GM dimension** predicted T2 proactive overt (b= .14; p<0.05) and reactive overt (b= 0.10; p<0.05) aggressive behavior. • T1 **II dimension** predicted T2 reactive overt (b= 0.09; p<0.05) aggressive behavior.

Note.
RAP = Rural Adaptation Project
COPAL= Constuyendo Oportunidades Para Adolecentes Latinos
ESOL= English for Speakers of Other Languages
CU = callous unemotional
GM = grandiose-manipulative
II = impulsive irresponsible

Jambroes et al., 2018; Lui et al., 2017). Finigan-Carr et al. (2015) investigated on how perceived behavioural control (self-control and decision making) influenced the overall framework of risk and protective factors for aggressive behaviour, whereas Lui et al. (2017), Barry et al. (2018), and Jambroes et al. (2018) examined the association between psychopathic traits (i.e., callous-unemotional traits, narcissism, impulsive irresponsible) and adolescent aggression. Other assessed individual factors which can be associated with adolescent aggression are emotional components such as anger, hostility (Rosenbaum et al., 2018), and lack of empathy (van Hazebroek et al., 2016). Social predictors investigated included bad peer influence (Lui et al., 2017; Wang et al., 2017) school climate (Wang et al., 2017), parental factor (Finigan-Carr et al., 2015), and substance abuse (Kivimaki et al., 2014).

As for the cohort studies, there were seven selected articles with sample size ranged from 302 (Forster et al., 2015) to 8,333 (Stalker et al., 2018) participants. All the studies were conducted in developed countries: United States (Forster et al., 2015; Reyes et al., 2015; Sittner & Hautala, 2016; Smokowski et al., 2016; Stalker et al., 2018) and Spain (Calvete et al., 2015; Orue et al., 2016). Nevertheless, most of the participants' characteristics comprised those who were in marginal and vulnerable group such as those who live in rural area of developed countries (Sittner & Hautala, 2016; Smokowski et al., 2016), lower education background (Reyes et al., 2015), and Hispanic immigrants (Forster et al., 2015). Self-reported continuous measures were primarily collected in all studies by the participants except for one study that categorised aggressive delinquency (Sittner & Hautala, 2016). Psychological predictors looked into including personality traits (Orue et al., 2016) and emotional components (Calvete et al., 2015). As for social predictors, peer influence (Smokowski et al., 2016; Calvete, 2015; Forster et al., 2015), school climate (Sittner & Hautala, 2016; Smokowski et al., 2016; Stalker et al., 2018), substance use (Reyes et al., 2015; Sittner & Hautala, 2016), and parent-adolescent conflict (Sittner & Hautala, 2016; Smokowski et al., 2016) were investigated in aggressive behaviour among adolescents. Apart from that, religion and cultural factors also are being observed among school students in rural area (Smokowski et al., 2016) and immigrants (Forster et al., 2015).

Predictors of Adolescent Aggression

The elicited adolescent aggression predictors from selected articles can be mainly divided into psychological predictors (Table 4) and social predictors (Table 5). Both consist of risk and protective factors.

The psychosocial predictors of adolescent aggression were described across different study designs of cross-sectional and cohort studies. Psychological predictors that were most highly reported by four articles were on the personality traits, mainly callous unemotional (Barry et al., 2018; Jambroes et al., 2018; Lui et al., 2017; Orue et al., 2016). Callous-

Table 4
Summary of psychological predictors of adolescent aggression

Factors	Predictors	Author/years	Main findings
Risk Factors	Personality traits	Barry et al., 2018; Jambroes et al., 2018; Lui et al., 2017; Orue et al., 2016	Callous unemotional, grandiose-manipulative, impulsive-irresponsible, narcissism
	Emotional factors	Calvete, 2015; Rosenbaum et al., 2018; Smokowski et al., 2016	Anger, hostility, internalizing symptoms
	Motive	van Hazebroek et al., 2017	Agentic goals (desire to be dominant)
Protective Factors	Self-control	Finigan-Carr et al., 2015; Rosenbaum et al., 2018	Self-control skills

Table 5
Summary of social predictors of adolescent aggression

Factors	Predictors	Author/years	Main findings
Risk Factors	Peer influence	Finigan-Carr et al., 2015; Forster et al., 2015; Sittner & Hautala, 2016; Smokowski et al., 2016; Wang et al., 2017	Problem with friend influence, delinquent peer association and affiliation, peer pressure, friend rejection, early dating
	Substance use	Kivimaki et al., 2014; Reyes et al., 2015; Sittner & Hautala, 2016	Ever tried substance, alcohol use, hard drug use
	Negative school climate	Smokowski et al., 2016; Stalker et al., 2018	School hassles
	Neighbourhood influence	Forster et al., 2015; Smokowski et al., 2016	Less educated residents, context of reception, openness/hostility
	Parent-adolescent conflict	Smokowski et al., 2016	Parental rejection
Protective Factors	Positive school climate	Sittner & Hautala, 2016; Smokowski et al., 2016; Wang et al., 2017	High quality school climate, positive school adjustment, school satisfaction
	Parent and Family factors	Smokowski et al., 2016; Wang et al., 2017	Parental disapproval of fighting, two-parent family structure
	Religion and culture	Smokowski et al., 2016	Religious orientation, ethnicity identity

unemotional dimension of psychopathy predicted reactive (β= .26; 95%CI= 0.02-0.36) and proactive aggression (β= 0.25; 95%CI= 0.02-0.36) among adolescents in Amsterdam (Jambroes et al., 2018), supported by other research evidence in different population (Barry et al., 2018; Orue et al., 2016). Other personality components of aggressive adolescents include grandiose-manipulative (Orue et al., 2016), impulsive-irresponsible (Jambroes et al., 2018; Orue et al., 2016), and narcissism (Barry et al., 2018). Emotional factors such as anger (β= 0.08, $p < 0.01$), hostility (β= 0.14; $p<0.001$), and internalizing symptoms (ExpB= 1.013, $p<0.001$) are also being reported by three articles (Calvete et al., 2015; Rosenbaum et al., 2018; Smokowski et al., 2016) to be significantly predicted adolescent aggression.

As part of social goal, agentic goals as one's desire to be dominant predicted aggressive behaviour among adolescents in single study conducted in Netherland (van Hazebroek et al., 2016). Despite the above-mentioned psychological risk factors, self-control among adolescents can prevent aggression development by empowering self-control skills ($\beta = -0.07$; $p<0.05$) as evidenced from two studies (Finigan-Carr et al., 2018; Rosenbaum et al., 2018).

According to the most commonly reported social predictors based on number of articles in this review, peer influence is the highest (five articles), followed by school climate (four articles), substance use (three articles), neighbourhood influence (two articles), family and parent factors (two articles), religion and culture factors (one article). All included studies' outcome measurements include physical and/or verbal aggression with sex-adjusted data analysis in predicting adolescent aggression.

The influence of friend with problematic behaviours like smoking and drinking alcohol as perceived norm factor was found to be a significant predictor of aggressive behaviours among those with medium (AOR= 2.61; 95%CI= 1.68-4.07) and high levels (AOR= 5.05; 95%CI= 3.03-48.39) of problem friend influence (Finigan-Carr et al., 2015). Selective affiliation of adolescents with peers who show serious problematic behaviour such as fighting also predicted aggressive behaviour among this age group (Wang et al., 2017). Early dating (RR= 2.65-2.76; $p<0.05$) and delinquent peer association (RR= 2.05-2.94; $p<0.05$) increased the risk of aggressive delinquency among North American Indigenous adolescents almost three times across different group-based trajectory (Sittner & Hautala, 2016). Delinquent peers are also being supported by Smokowski et al. (2016) (ExpB= 1.124; $p<0.001$) and Forster et al. (2015) ($\beta= 0.1024$; $p=0.047$). Additionally, friend rejection (ExpB= 1.023; $p<0.001$) and peer pressure (ExpB= 1.018; $p<0.01$) also predicted aggression in disadvantaged adolescents (Smokowski et al., 2016).

Ever tried substances (tobacco, alcohol, and marijuana) can be as high as almost five times likely among the chronic aggressive adolescents (RR= 4.78; $p<0.001$) and almost three times likely among moderate desistor, those who stopped from offending actions (RR=2.45; $p<0.05$), in trajectory groups of aggressive delinquency (Sittner & Hautala, 2016). Other substances used which can predict adolescent aggression in physical dating violence include heavy alcohol use ($\beta= 0.17$; $p<0.001$) and hard drug use ($\beta=0.05$; $p= 0.01$) such as cocaine, heroin, and ecstasy as described by Reyes et al. in 2015. Alcohol use and its level of consumption are also found to be associated with aggressive behaviour among both male and female school students in Finland (Kivimaki et al., 2014).

School climate and parent factors can be both promotive and protective factors. School hassles as proxy to school climate were shown to predict adolescent aggression in several studies (Smokowski et al., 2016; Stalker et al., 2018). On the other hand, higher quality of school climate by teacher, student-student support, and opportunities for autonomy in school factor had negatively predicted aggression among Chinese adolescents (Wang et al., 2017). Positive school satisfaction (Smokowski et al., 2016) and adjustment (Sittner & Hautala, 2016) also act as protective factors against adolescent aggression development. As for parental factors, parent-adolescent conflict predicted adolescent aggression (ExpB= 1.013, $p<0.001$) as described by Smokowski et al. (2016). However, parental disapproval of fighting (Finigan-Carr et al., 2015) and two-parent family structure (Smokowski et al., 2016) prevented aggressive behaviour development among school students in United States.

Other minimal social predictors of adolescent aggression include neighbourhood influence of less educated residents (Smokowski et al., 2016). However, this negative predictor can be buffered by protective factors of religious orientation and ethnicity identity. As for the development of aggressive behaviour among new immigrant adolescents in United States (Forster et al., 2015), context of reception by means of social supports and economic opportunities and openness or hostility by the local community also play significant role ($\beta= 0.184$; $p=0.021$).

DISCUSSION

This systematic review assessed on 15 observational studies (cross-sectional and cohort) in determining adolescent aggression predictors. Though causal relationship effect can be described more effectively from cohort study, such association can also be predicted in cross-sectional studies with appropriate statistical analysis application. The highly reported psychological predictors of adolescent aggression elicited from selected articles were personality traits, emotional factors, and self-control. As for social predictors, peer influence factor was the most frequently stated in affecting aggressive behaviour among adolescents. This is followed by substance use, school climate, parent and neighbourhood influence.

Each study article involved different adolescent with the age range of 10 to19 years old, which comprised of middle/secondary school students. As the age of onset and peak aggression among adolescents can vary in their subtypes and nature (Clow, 2016), the different adolescent age range in the selected study articles may somehow affect the findings of the review. For example, delinquent peer associations can be a significant risk factor for aggressive behaviour among chronic desistor, who started off at a moderate level of aggression, increased and peaked at a high level of aggression at approximately 15 years of age, and decreased in later adolescence. As for adolescent-limited offender who started off at low level of aggression and peaked at moderate level of aggression before subsided, this factor can be a significant protective factor when compared to those who were chronic desistors (Sittner & Hautala, 2016).

Callous unemotional personality trait was the highest psychological predictor of adolescent aggression from the systematic review. The callousness (inflicting harm and lack of empathy) and uncaring (general disregards for rules, performance or norms) dimensions of CU traits are the most strongly linked to aggression (Berg et al., 2013; Roose et al., 2010). As these two dimensions are connected to both hostility and general approval of aggression (Lui et al., 2017), CU traits are also exceptionally related to more severe types of aggressive behaviour of proactive overt and proactive relational (Orue et al., 2016). CU traits are also found to have stronger prediction for proactive aggression in adolescents with high verbal-intelligence scores (Jambroes et al., 2018), which can be more established in adolescents as compared to younger children.

Social predictors were found to be more common and higher in the strength of evidence presented as compared to psychological predictors from the review. Bad influence from problematic friend and ever trying substances were among the highest predictors of adolescent aggression, which could increase the risk by three to five times more likely compared to those non-offenders (Sittner & Hautala, 2016). Early antisocial influences among adolescents might increase their opportunities to learn pro-delinquent beliefs (Haynie & Osgood, 2005) as most of them spent their entire childhood and adolescence within the same peer groups (Whitbeck et al., 2014). As for the result, peer influence was one of the salient predictors in adolescents' behavioural development, which can be difficult to control.

Negative proximal processes of school, parents, and neighbourhood predictors were found to be minimally affected by the development of aggressive behaviour with less than 10% of risk (Smokowski et al., 2016). One of the reasons could be due to the drawback of the study articles that did not differentiate between proactive and reactive aggression or between direct (overt) and indirect (relational) aggression. Adolescents' response towards their social environment might vary based on the context and situations, which could also influence the purpose of their actions. Furthermore, the relationship of social predictors such as family violence and neighbourhood social control with aggressive behaviour among adolescents might not be directly associated but as moderating effect (Reyes et al., 2015). Findings from local evidences in Malaysia, reported some researches showed on family problem environment and functioning influenced on adolescent aggression (Azmawati et al., 2015; Zainah et al., 2011), while some researches showed that individual factors including self-esteem (Arokiaraj et al., 2011) and increase desire to try new things and freedom of action (Sharif & Roslan, 2011) played more important role as compared to family environment.

The limitations of reviewed articles that were selected can be derived from the quality of research methodology, which resulted in some risk of selection and information bias. Most of the researchers did not reveal the sample size justification, which might affect the statistical power of the study. The statistical methods that were used in the included studies were also heterogenous, with lack of effect estimates and 95% confidence interval information. Hence, meta-analysis conducted might not be feasible in this review. As for the limitations of search strategy of this review, only five-recent years published journal articles in English were included. The initial search process and articles screening were conducted by single author within limited time frame, which may result in selection bias during data extraction and possibility of related articles not being included in the final review.

Based on the United States Preventive Service Task Force guidelines, the strength of evidence for the selected observational studies were moderate. The available evidence was sufficient to determine the predictors of aggression but constrained by main weaknesses of non-generalized study population, social-desirability bias from self-reporting, and limited participants' information for potentially important variables of interest. Though most of the study participants were school students, some articles focused on school drop-out adolescents with problem misconduct, who received military programmes and closed treatment closed treatment setting for adolescents with severe conduct problems (Lui et al., 2017; Jambroes et al., 2018). Their age ranges were older compared to other study participants (15 to 19 years old), which was assumed to have more developed and stable socio-cognitive processes compared to younger sample (Lui et al., 2017). Findings among these school drop-outs also stressed on the individual factors including callous unemotional personality traits (Lui et al., 2017) and impulsive irresponsible psychopathy (Jambroes et al., 2018). However, their findings may not be generalized to those attending public schools as they were mainly selected by purposive sampling and might differ in their previous beliefs and perceptions in regards to hostility and approval of aggression (Lui et al., 2017). Findings from several articles which looked into specific ethnicity groups in rural region (Sittner & Hautala, 2016) and vulnerable groups of immigrants (Forster et al., 2015) and undersiege people (Rosenbaum et al., 2018), also limit their generazibility of findings to general population of adolescents.

As for self-report measure, it has potential bias due to same source, selective memory and social desirability of respondents (Forster et al., 2015; Lui et al., 2017; Reyes et al., 2015). They also might be influenced by the presence of their peers while responding to the questionnaires as they were conducted during school sessions (Smokowski et al., 2016). However, evidences had suggested that self-report was better compared to peer-reports in determining aggressive behaviour predictors (Little et al., 2003) and yielded reliable results (Bradburn et al., 1987; Rutherford et al., 2000).

Psychosocial treatments were shown to be effective in aggressive adolescents with conduct disorder (Brestan & Eyberg, 1998), which might include problem-solving skills training, parent management training, functional family therapy and multisystemic therapy (Kazdin, 1997). The retrieved psychosocial predictors of adolescent aggression from this review can be important information for the stakeholders namely parents, educators, psychologists, and policy makers in their efforts to reduce the social misconduct in the community. This warrant for more practical and holistic approach in dealing with aggressive adolescents by incorporating the important psychosocial predictors so they can be applied in more targeted and updated intervention strategy deliveries publicly.

More researches that further explore these significant predictors could also provide further in-depth understanding to explain this complex and myriad processes of adolescent aggression development. Future studies may also incorporate various data sources in investigating adolescents' aggressive behaviour by cross-validating self-reports with parent, teacher or peer reports, especially for psychopathic traits that can be most susceptible to socially desirable response (Lui et al., 2017; Rosenbaum et al., 2018).

CONCLUSION

From the systematic review of related articles, the main psychosocial predictors of adolescent aggression include personality traits, emotional factors, peer influence, and substance abuse, which all need to be focused in preventive strategies against adolescent aggression.

ACKNOWLEDGMENTS

We would like to thank the Director General of Health Malaysia for his permission to publish this article. Special thanks to Professor Dr. Zamberi Sekawi, the Dean of Faculty of Medicine and Health Sciences, Universiti Putra Malaysia, for allowing us to publish this paper. High appreciation also to Universiti Putra Malaysia for providing Putra Grant - Putra Graduate Initiative (GP-IPS) for the main author's research conduct related to this manuscript.

REFERENCES

Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for ASEBA school-age forms and profiles.* Burlington, VT: University of Vermont, Research Center for Children, Youth & Families.

Arokiaraj, A. S., Nasir, R., & Shahrazad, W. S. W. (2011). Gender effects on self-esteem, family functioning and resilience among juvenile delinquents in Malaysia. *Pertanika Journal of Social Sciences and Humanities, 19*(S), 1-8.

Azmawati, M. N., Hamid, A., Hazariah, S., Shamsul, A. S., Norfazilah, A., Azimatun, N. A., & Rozita, H. (2015). Risk taking behaviour among urban and rural adolescents in two selected districts in Malaysia Risk taking behaviour among urban and rural adolescents in two selected districts in Malaysia. *South African Family Practice*, *57*(3), 160-165.

Barry, C. T., McDougall, K. H., Anderson, A. C., & Bindon, A. L. (2018). Global and contingent self-esteem as moderators in the relations between adolescent narcissism, callous-unemotional traits, and aggression. *Personality and Individual Differences*, *123*, 1-5.

Berg, J. M., Lilienfeld, S. O., Reddy, S. D., Latzman, R. D., Roose, A., Craighead, L. W., ... & Raison, C. L. (2013). The inventory of callous and unemotional traits: A construct-validational analysis in an at-risk sample. *Assessment*, *20*(5), 532-544.

Bosworth, K., & Espelage, D. (1995). *Teen conflict survey*. Bloomington, Indiana: Centre for Adolescent Studies, Indiana University.

Bradburn, N., Rips, L., & Shevell, S. (1987). Answering auto-biographical questions: The impact of memory and inference on surveys. *Science, 235*(4798), 157-161.

Brestan, E. V., & Eyberg, S. M. (1998). Effective psychosocial treatments of conduct- disordered children and adolescents: 29 years, 82 studies, and 5272 kids. *Journal of Clinical Child Psychology*, *27*(2), 180-189.

Buss, A. H., & Perry, M. (1992). The aggression questionnaire. *Journal of Personality and Social Psychology 63*(3), 452-459.

Butchart, A., & Mikton, C. (2014). *Global status report on violence prevention, 2014*. Geneva, Switzerland: World Health Organization publications.

Calvete, E., Gamez-Guadix, M., & Garcia-Salvador, S. (2015). Social information processing in child-to-parent aggression: Bidirectional associations in a 1-year prospective study. *Journal of Child and Family Studies*, *24*(8), 2204-2216.

Calvete, E., Gámez-Guadix, M., Orue, I., González-Diez, Z., Lopez de Arroyabe, E., Sampedro, R., ... & Borrajo, E. (2013). Brief report: the adolescent child-to parent aggression questionnaire: An examination of aggression against parents in Spanish. *Journal of Adolescence, 36*, 1077-1081.

Card, N. A., & Little, T. D. (2006). Proactive and reactive aggression in childhood and adolescence: A meta-analysis of differential relations with psychosocial adjustment. *International Journal of Behavioral Development, 30*, 466-480.

Carubia, B., Becker, A., & Levine, B. H. (2016). Child psychiatric emergencies: updates on trends, clinical care, and practice challenges. *Current Psychiatry Reports*, *18*(4), 41-48.

Cima, M., Raine, A., Meesters, C., & Popma, A. (2013). Validation of the Dutch reactive proactive questionnaire (RPQ): differential correlates of reactive and proactive aggression from childhood to adulthood. *Aggressive Behavior*, *39*(2), 99-113.

Clow, K. A. (2016). *Aggression among children and youth: an examination of service allocation for subtypes of aggression* (Master thesis). University of Western Ontario, Canada.

Conner, D. F., & Barkley, R. A. (2004). *Aggression and antisocial behavior in children and adolescents: research and treatment*. New York, NY: The Guilford Press.

Duru, C. K. (2015). *Antisocial personality and aggressive behaviour intention among secondary school children in Malaysia* (Unpublished Doctoral thesis). Universiti Putra Malaysia, Malaysia.

Finigan-Carr, N. M., Cheng, T. L., Gielen, A., Haynie, D. L., & Simons-Morton, B. (2015). Using the theory of planned behavior to predict aggression and weapons carrying in urban African American early adolescent youth. *Health Education and Behavior*, *42*(2), 220-230.

Forster, M., Grigsby, T., Soto, D. W., Schwartz, S. J., & Unger, J. B. (2015). The role of bicultural stress and perceived context of reception in the expression of aggression and rule breaking behaviors among recent-immigrant Hispanic youth. *Journal of Interpersonal Violence*, *30*(11), 1807-1827.

Hamid, S. R. A., Ismail, K., Saad, S., Ibrahim, M. B., & Mansur, N. (2015). Are the Malaysian adolescents' behavior at-stake? *Mediterranean Journal of Social Sciences*, *6*(2 S5), 328-346.

Haynie, D. L., & Osgood, D. W. (2005). Reconsidering peers and delinquency: how do peers matter? *Social Forces, 84*, 1109-1130.

Hussin, S. F. M., Aziz, N. S. A., Hasim, H., & Sahril, N. (2014). Prevalence and factors associated with physical fighting among Malaysian adolescents. *Asia Pacific Journal of Public Health, 26*(5_suppl), 108S-115S.

Ibu Pejabat Polis Kontijen. (2018). *Perangkaan penglibatan pelajar dalam jenayah indeks bagi bulan Jan-Dis 2015-2017* [Statistics on student involvement in index crime for month of Jan-Dis 2015-2017]. Shah Alam, Malaysia: Bahagian Siasatan Jenayah Kontijen Selangor.

Jambroes, T., Jansen, L. M., vd Ven, P. M., Claassen, T., Glennon, J. C., Vermeiren, R. R., ... & Popma, A. (2018). Dimensions of psychopathy in relation to proactive and reactive aggression: does intelligence matter? *Personality and Individual Differences*, *129*, 76-82.

Kazdin, A. E. (1997). Practitioner review: Psychosocial treatments for conduct disorder in children. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *38*(2), 161-178.

Kim, S., Kamphaus, R. W., Orpinas, P., & Kelder, S. H. (2010). Change in the manifestation of overt aggression during early adolescence: Gender and ethnicity. *School Psychology International, 31*(1), 95-111.

Kivimaki, P., Kekkonen, V., Valtonen, H., Tolmunen, T., Honkalampi, K., Tacke, U., … & Laukkanen, E. (2014). Alcohol use among adolescents, aggressive behaviour, and internalizing problems. *Journal of Adolescence, 37*, 945-951.

Krug, E., Dahlberg, L., Mercy, J., Zwi, A., & Lozano, R. (2002). *World Health Organization: World report on violence and health.* Geneva, Switzerland: World Health Organization publications.

Little, T., Jones, S., Henrich, C., & Hawley, P. (2003). Disentangling the "whys" from the "whats" of aggressive behaviour. *International Journal of Behavior Development, 27*, 122-133.

Liu, J. (2004). Prenatal and perinatal complications as predispositions to externalizing behavior. *Journal of Prenatal and Perinatal Psychology and Health*, *18*(4), 301-312.

Liu, J., Lewis, G., & Evans, L. (2013). Understanding aggressive behaviour across the lifespan. *Journal of Psychiatric and Mental Health Nursing*, *20*(2), 156-168.

Lui, J. H., Barry, C. T., & Schoessler, M. (2017). The indirect effects of adolescent psychopathic traits on aggression through social-cognitive factors. *Journal of Child and Family Studies*, *26*(5), 1298-1309.

Macleod, J., & Smith, G. D. (2003). Psychosocial factors and public health: a suitable case for treatment? *Journal of Epidemiology and Community Health*, *57*(8), 565-570.

Marsee, M. A., Barry, C. T., Childs, K. K., Frick, P. J., Kimonis, E. R., Munoz, L. C., ... & Lau, K. S. L. (2011). Assessing the forms and functions of aggression using self-report: Factor structure and invariance of the peer conflict scale in youths. *Psychological Assessment, 23*, 792-804.

Martikainen, P., Bartley, M., & Lahelma, E. (2002). Psychosocial determinants of health in social epidemiology. *International Journal of Epidemiology*, *31*(6), 1091-1093.

Masoumeh, H. (2014). *Relationships between emotional intelligence, parent-child relationship, deviant peer affiliation, and aggressive behavior among adolescents in Tehran, Iran* (Unpublished Doctoral thesis). Universiti Putra Malaysia, Malaysia.

Maxwell, J. P. (2008). Psychometric properties of a Chinese version of the Buss-Warren Aggression Questionnaire. *Personality and Individual Differences, 44*, 943-953.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D.G. (2009). Preferred reporting items for systematic reviews and Meta-analyses: the PRISMA statement. *Annals of Internal Medicine, 151*(4), 264-269.

Orue, I., Calvete, E., & Gamez-Guadix, M. (2016). Gender moderates the association between psychopathic traits and aggressive behavior in adolescents. *Personality and Individual Differences, 94*, 266-271.

Polman, H., de Castro, B. O., Thomaes, S., & van Aken, M. (2009). New directions in measuring reactive and proactive aggression: validation of a teacher questionnaire. *Journal of Abnormal Child Psychology, 37*(2), 183-193.

Psychosocial. (2019). *Oxford Living Dictionaries*. Retrieved May 30, 2018, from https://en.oxforddictionaries.com/definition/psychosocial

Rashtriya, K., & Swasthya, K. (2014). *Training module for ANMs/LHVs to provide Adolescent-Friendly Health Services*. New Delhi, India: Ministry of Health and Family Welfare.

Reyes, H. L. M., Foshee, V. A., Tharp, A. T., Ennett, S. T., & Bauer, D. J. (2015). Substance use and physical dating violence: the role of contextual moderators. *American Journal of Preventive Medicine*, *49*(3), 467-475.

Roose, A., Bijttbier, P., Decoene, S., Claes, L., & Frick, P. J. (2010). Assessing the affective features of psychopathy in adolescence: A further validation of the Inventory of callous and unemotional traits. *Assessment, 17*, 44-57.

Rosenbaum, M., Ronen, T., Abuelaish, I., Orkibi, H., & Hamama, L. (2018). Positivity ratio links self-control skills to physical aggression and happiness in young Palestinians living in Gaza. *Young*, *26*(4_suppl), 78S-95S.

Rutherford, M. J., Cacciola, J. S., Alterman, A. I., McKay, J. R., & Cook, T. G. (2000). Contrasts between admitters and deniers of drug use. *Journal of Substance Abuse Treatment, 18*(4), 343-348.

Seong, L. C. (2008). *Relationships between parents' marital quality, family environment and students' behaviour of selected secondary schools in Selangor and Kuala Lumpur, Malaysia* (Unpublished Doctoral thesis). Universiti Putra Malaysia, Malaysia.

Shaffer, D., Fisher, P., Lucas, C. P., Dulcan, M. K., & Schwab-Stone, M. E. (2000). NIMH diagnostic interview schedule for children version iv (NIMH DISC-IV): description, differences from previous versions, and reliability of some common diagnoses*. Journal of the American Academy of Child and Adolescent Psychiatry, 39*, 28-38.

Sharif, Z., & Roslan, M. N. (2011). Faktor-faktor yang mempengaruhi remaja terlibat dalam masalah sosial di Sekolah Tunas Bakti, Sungai Lereh, Melaka [Factors influencing adolescents to involve in social problems in Tunas Bakti School, Sungai Lereh, Melaka]. *Journal of Education Psychology and Counseling*, *1*(7), 115-140.

Sittner, K. J., & Hautala, D. (2016). Aggressive delinquency among north American indigenous adolescents: trajectories and predictors. *Aggressive Behavior*, *42*(3), 274-286.

Smokowski, P. R., Guo, S., Cotter, K. L., Evans, C. B., & Rose, R. A. (2016). Multi-level risk factors and developmental assets associated with aggressive behavior in disadvantaged adolescents. *Aggressive Behavior*, *42*(3), 222-238.

Srinath, S., Kandasamy, P., & Golhar, T. S. (2010). Epidemiology of child and adolescent mental health disorders in Asia. *Current opinion in psychiatry, 23*(4), 330-336.

Stalker, K. C., Wu, Q., Evans, C. B., & Smokowski, P. R. (2018). The impact of the positive action program on substance use, aggression, and psychological functioning: Is school climate a mechanism of change? *Children and Youth Services Review*, *84*, 143-151.

van Hazebroek, B., Olthof, T., & Goossens, F. A. (2016). Predicting aggression in adolescence: The interrelation between (a lack of) empathy and social goals. *Aggressive Behavior*, *43*(2), 204-214.

Wang, Z., Yu, C., Zhang, W., Chen, Y., Zhu, J., & Liu, Q. (2017). School climate and adolescent aggression: A moderated mediation model involving deviant peer affiliation and sensation seeking. *Personality and Individual Differences*, *119*, 301-306.

Whitbeck, L. B., Hartshorn, K., & Walls, M. (2014). *Indigenous adolescent development: psychological, social and historical contexts*. New York, NY: Routledge.

World Health Organization. (2006). *National report on violence and health-Malaysia.* Kobe, Japan: Center for Health Development.

World Health Organization. (2017). *Adolescents: health risks and solutions*. [Factsheet]. Retrieved May 30, 2018, from http://www.who.int/mediacentre/factsheets/fs345/en/

Yang, Y., Joshi, S. H., Jahanshad, N., Thompson, P. M., & Baker, L. A. (2016). Neural correlates of proactive and reactive aggression in adolescent twins. *Aggressive Behavior*, *43*(3), 230-240.

Zainah, A. Z., Nasir, R., Khairudin, R., & Shahrazad, W. S. W. (2011). Family functioning as predictor towards self-concept among delinquent adolescents. *Pertanika Journal of Social Sciences and Humanities, 19*(S), 83-89.

Zirpoli, T. J. (2008). *Behavior management: Applications for teachers*. New Jersey, USA: Pearson/ Merrill Prentice Hall.

Pertanika J. Sci. & Technol. 27 (4): 1509 - 1525 (2019)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Review article

Timing and Prognostic Factors of Tuberculosis Treatment Interruption

Qudsiah Suliman[1,2], Salmiah Md. Said[1]* and Lim Poh Ying[1]

[1]*Department of Community Health, Faculty of Medicine and Health Sciences, Universiti Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia*
[2]*Ministry of Health, Federal Government Administration Centre, 62590, Putrajaya, Malaysia*

ABSTRACT

The global Tuberculosis epidemic (TB) poses a significant public health threat. While the consequences of TB treatment interruption are indisputable, the knowledge about the timing and prognostic factors of TB treatment interruption is fundamental. Despite a considerable amount of evaluation, the timing and prognostic factors of TB treatment interruption have been inconsistently identified from one study to another. Therefore, this study aimed to examine the evidence obtained from published literature on the timing and prognostic factors of TB treatment interruption at different points of the treatment course. In this review, three databases namely *Pubmed, Scopus*, and *Science Direct* were used to identify articles published from January 2003 to February 2018. This was based on the inclusion criteria and keywords including 'default', 'survival time', 'tuberculosis', and 'treatment interruption'. The nine selected studies were prospective and retrospective cohort studies conducted in developing countries. The diversity of the study's participants and TB treatment interruption definition were allowed, thus delineating a heterogeneous finding. This review suggests that the interruption predominantly occurred during the maintenance phase of treatment course. Despite the finding, a considerable gap in understanding the prognostic factors at different time points of TB treatment interruption was elicited. The heterogeneity across the studies may limit the inferences and warrant further evaluation. In essence,

ARTICLE INFO

Article history:
Received: 02 July 2018
Accepted: 13 May 2019
Published: 21 October 2019

E-mail addresses:
qudsiahsuliman@yahoo.com (Qudsiah Suliman)
salmiahms@upm.edu.my (Salmiah Md. Said)
poh_ying@upm.edu.my (Lim Poh Ying)
* Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

the time-related information should be integrated into framing impactful public health strategy, while a vigorous attempt on the evaluation of the cognitive, behavioural and psychosocial aspects may be beneficial.

Keywords: Default, survival time, treatment interruption, tuberculosis

INTRODUCTION

The 20th century has demonstrated a significant public health threat posed by the re-emergence of an ancient airborne disease which is Tuberculosis (TB). The TB mortality has been substantial, reaching approximate 1.6 million in 2016, while the slow falling rate of TB incidence has been illustrated across the nations (World Health Organization, 2017). This has called for an intensified global effort to end TB global epidemic via End TB strategy. As one of the major obstacles in successfully treating 85% of the detected TB cases, TB treatment interruption should be the focus area of improvement in TB management (World Health Organization, 2017).

TB treatment interruption is defined as a history of stopping treatment for two or more consecutive months (World Health Organization, 2014). Evidently, TB treatment interruption has led to devastating consequences including delayed sputum conversion, drug resistance, longer treatment regimes, incomplete treatment, hence collectively exhibit prolonged infectiousness to the community (Dominguez-Castellano et al., 2003; Kuaban et al., 2009; Marx et al., 2012; Nahid et al., 2011; Pablos-Méndez et al. 1996; Vree et al., 2007). Importantly, TB treatment interruption has also posed significant economic burden and dependency through increased hospital admission and bed occupancy, thus leading to significant psychosocial impact (Cerdá et al., 2014; Pettit et al., 2013).

The recent two decades have demonstrated vigorous efforts to combat TB treatment interruption. In parallel with the defaulter tracing system, Directly Observed Treatment, Short-course (DOTS) has been a global core strategy to increase treatment adherence since 1993 (Maher, 2009; World Health Organization, 2017). On top of fee exemption for the anti-TB drug in most developing countries, fixed-dose combination tablets for tuberculosis treatment have been advocated since the 1980s, primarily to ensure optimal adherence (World Health Organization, 1999). Notwithstanding the established policy and legislative framework intended to optimise acceptance and access to treatment, TB treatment interruption remains a global public health challenge across nations particularly in high disease burden countries (Sabate, 2003; World Health Organization, 2017). In developing and high TB burden countries, TB treatment interruption has been reported ranging from 6 to 30% (Connolly et al., 1999; Dodor, 2004; Kliiman & Altraja, 2010).

Previous studies delineated risk factors of TB treatment interruption. These include younger age, male, low educational level, unemployment, being immigrant and HIV infection, which had significantly influenced TB treatment default (Chee et al., 2000;

Connolly et al., 1999; da Silva Garrido et al., 2012; Kliiman & Altraja, 2010; Millet et al., 2009; Naing et al., 2001; Tachfouti et al., 2013). Meanwhile, smoking and alcoholism have been consistently reported as the risk factors of TB treatment interruption (Bagchi et al., 2010; Chandrasekaran et al., 2005; Kliiman & Altraja, 2010; Sulaiman, & Ali, 2010; Vijay et al., 2010).

In the recent years, the outstandingly high burden due to TB treatment interruption has resulted in major awareness about the importance of assessing the timing of interruptions and determinants of the interruptions at different stages of the treatment course (Kruk et al., 2008). While much has been elicited on the determinants of TB treatment interruption, to the best of our knowledge there has been a scarce systematic review on the timing of interruption and prognostic factors contributing to different timing of treatment interruption. In addition, there are a number of prognostic factors reported previously, but few had been consistently identified from one study to another (Jenkins et al., 2013; Jepchumba et al., 2017; Akessa et al., 2015). Therefore, this review aims to examine evidence from published literature on the timing of TB treatment interruption in TB treatment, and subsequently to assess prognostic factors of TB treatment interruption at different time points of the treatment course. The duration information from survival analysis is fundamental to assist in designing time relevant intervention strategies in TB case holding and management, at which a different timeline of impactful adherence strategy can be planned accordingly.

METHODS

Data Sources and Literature Search Strategy

Literature search was systematically conducted via three electronic databases, namely *PubMed, Scopus*, and *ScienceDirect* in February 2018. The *Pubmed* search terms were (“survival”[MeSHTerms] OR (“survival”[All Fields] OR “survival time”[MeSHTerms]) OR (“survival time”[AllFields]) OR (“time*”[MesHTerms]) OR (“time*”[AllFields]) AND (“tuberculosis”[MeSHTerms] OR “tuberculosis”[AllFields]) AND (“treatment interruption”[MeSHTerms] OR “treatment interruption”[AllFields] OR (“treatment default”[MeSHTerms]OR “treatment default”[AllFields] OR (“default”[MeSHTerms]OR “default”[AllFields]). A similar search strategy was adapted in order to search for articles from *ScienceDirect*, and *Scopus*. The search for scholarly literature was also extended to *Google Scholar*.

All retrieved abstracts and titles were read and screened by three independent investigators (Q.S, S.M.S, and L.P.Y). The inclusion criteria were; original studies in English language, cohort studies of patients with tuberculosis, studies with TB treatment interruption reported as the primary outcome, as well as papers presenting any temporal data on TB treatment interruption such as mean/median time to default, survival probability, cumulative hazard or hazard function. In terms of TB treatment interruption, a variety of criteria for defining interruption was accepted (for example early treatment interruption, or interruption of two or more months). Due to the scarcity of relevant articles, the date range was expanded to include published articles in peer-reviewed journals from February 2003 to February 2018, instead of limiting the search only to the past 10 years. The types of papers excluded from the analysis were reviews or editorials, non-peer-reviewed review literature such as technical reports and web-based guidelines, articles focusing on unfavourable outcomes rather than TB treatment interruption, as well as articles describing temporal data without the assessment of prognostic factors. Full text of potentially eligible articles were retrieved and screened for eligibility, through which the discrepancies between the reviewers' preferences on potentially eligible articles were resolved.

Search Outcome

The search was electronically conducted according to the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) guidelines, 2009 (Moher et al., 2009). The PRISMA flow diagram for study selection is described in Figure 1. Firstly, 267 articles

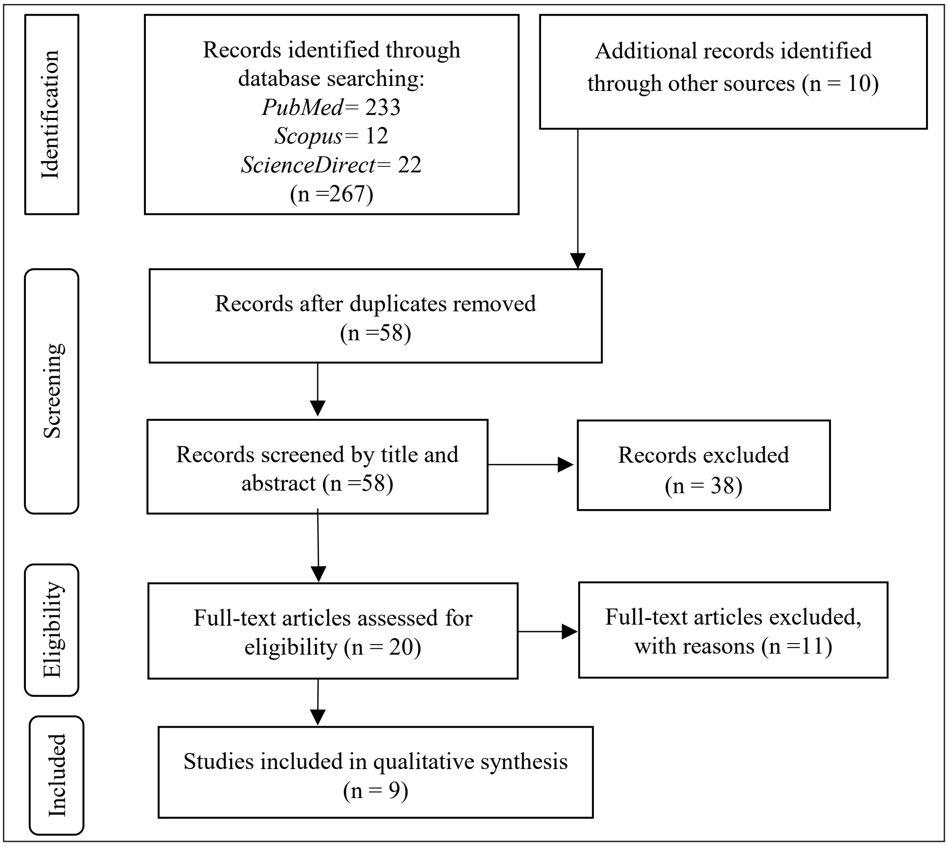


Figure 1. PRISMA flowchart for systematic search on timing and prognostic factors of TB treatment interruption

were identified from three databases. After excluding duplicated articles, assessment of titles and abstracts was executed, through which unrelated articles were omitted from further screening. After screening and looking for eligible articles based on the inclusion criteria, only 20 articles were examined. A sum of 11 articles were excluded due to the following reasons; articles focusing on unfavourable outcomes rather than TB treatment interruption or articles focusing on the survival time of treatment interruption without emphasizing prognostic factors.

Quality Assessment, Data Extraction and Data Analysis

The title, abstract and full text of every article retrieved from the search were screened by one reviewer (Q.S). To ensure rigour, a second set of reviewers (S.M.S and L.P.Y) were employed to assess the retrieved study for any doubt on the article's eligibility. The related articles were reviewed based on characteristics of the individual study including study details, exposure assessment, operational definition of treatment interruption, statistical method, as well as study outcomes including treatment interruption rate, survival time, survival probability and prognostic factors of TB treatment interruption. The included articles were evaluated for its risk of bias using Quality in Prognosis Studies (QUIPS) tool (Hayden et al., 2008). Assessment of bias was conducted based on six domains including the study's participants, study's attrition, prognostic factor measurement, outcome measurement, study confounding as well as statistical analysis and reporting. Some studies showed the lack of descriptions for prognostic factor measurement and study confounding particularly on the definition of prognostic factors, unclear methods for imputation of missing prognostic factors, as well as unclear description and measurement of important confounders. The list of studies is summarized in Table 1. Despite an overt criticism on the studies' quality and generalizability, it was not intended for article exclusion. Meanwhile, a meta-analysis of comparable studies was not conducted due to the presence of heterogeneity across the studies. Hence, a qualitative synthesis was performed.

RESULT

Finally, nine (9) eligible study articles were found to have met the predetermined study objectives and inclusion criteria. These articles were qualitatively synthesized, and the results were presented as follows.

Characteristic of Prognostic Studies

To date, several studies had been conducted to examine the time of TB treatment interruption and its prognostic factors. Most of the studies were designed as prospective cohort studies, and the majority of assessments were conducted in high TB burden countries. There had been a variation in the number of participants ranging from 249 to 90170 of patients

Table 1
Critical appraisal of prognostic studies on TB treatment interruption using quality in prognosis studies (QUIPS) tool

Study	Akessa et al. (2015)	Hill et al. (2005)	Jenkins et al. (2013)	Jepchumba et al. (2017)	Masini et al. (2016)	Pefura-Yone et al., (2014)	Rutherford et al. (2013)	Shargie & Lindtj (2007)	Sylvère (2015)
1. Study Participation: The study sample adequately represents the population of interest.									
Risk of bias	Low	Low	Moderate	Low	Low	Low	Low	Low	Moderate
2. Study Attrition: The study data available (i.e.,participants not lost to follow-up) adequately represent the study sample.									
Risk of bias	Low	Low	Moderate	Low	Low	Low	Low	Low	Moderate
3. Prognostic Factor Measurement: The PF is measured in a similar way for all participants.									
Risk of bias	Moderate	Low	Low	Low	Low	Low	Moderate	Low	Moderate
4. Outcome Measurement: The outcome of interest is measured in a similar way for all participants.									
Risk of bias	Low	Low	Low	Low	Low	Low	Low	Low	Low
5. Study Confounding: Important potential confounding factors are appropriately accounted for.									
Risk of bias	Moderate	Moderate	Low	Moderate	Low	Low	Moderate	Moderate	Low
6. Statistical Analysis and Reporting: The statistical analysis is appropriate, and all primary outcomes are reported.									
Risk of bias	Low	Low	Low	Low	Low	Low	Low	Low	Low

seeking treatments from both hospitals and health clinics. Most of the participants were from the middle-age group. In term of exposure assessment, there had been prominent diversity across the studies. Most of the studies evaluated the demography of the patients, organizational factors as well as biomedical assessments such as comorbidities and treatment-related factors, but a limited assessment was made in cognitive, behavioural and psychosocial aspects. The characteristics of the individual study are described in Table 2.

With the exemption of two studies conducted by Hill et al. (2005) and Masini et al. (2016), most of the studies had computed Cox Proportional Hazard in analysing time to event data. Due to violated proportionality assumption for the Cox regression, Hill et al. (2005) computed the Extended Cox Regression in his assessment. Meanwhile, Masini et al. (2016) incorporated both fixed effects and accounted for the geographic region of Kenya via a random effect's component, hence mixed effect Cox proportional hazard modelling was computed.

TB Treatment Interruption

The burden of TB treatment interruption for the individual study is summarized in Table 3. In general, the proportion of TB treatment interruption ranged from 5.0 to 25.2%. The three highest treatment interruption burdens were recorded according to assessments across high TB burden countries (Hill et al., 2005; Pefura-Yone et al., 2014; Shargie & Lindtj, 2007). In addition, there was considerable variation in term of the definition of TB treatment interruption used in the studies. Instead of defining TB treatment interruption as two consecutive months of interruption, Rutherford et al. (2013) and Hill et al. (2005) used different operational definitions of interruption which were two consecutive weeks and three consecutive days of interruption respectively.

Survival Time, Survival Probability or Cumulative Hazard of TB Treatment Interruption

In this review, median survival time refers to length of time from either the date of diagnosis or treatment initiation, after which 50% of the TB patients are still in the treatment course, or have yet to develop TB treatment interruption (Kleinbaum, 1996). Meanwhile, survival probabilities or survival function is the probability that the individual survives from the date of diagnosis or treatment initiation, until the last day of treatment. To date, the median survival time and survival probability had been differently reported across the study locations, as presented in Table 3. Most of the studies revealed that the median time of TB treatment interruption was depicted during the maintenance phase of TB treatment (Akessa et al., 2015; Hill et al., 2005; Jenkins et al., 2013; Pefura-Yone et al., 2014). As opposed to the preceding findings, a prospective cohort study involving 264 of newly diagnosed PTB smear-positive patients in Bandung, Indonesia revealed that the median time to treatment

Table 2
Characteristics of studies included in systematic review

Author/Year	Study Design /Settings/ Study location	Participants	Exposure Assessment	Operational Definition of Treatment Interruption
Akessa et al. (2015)	Retrospective cohort, Jimma University Hospital, Ethiopia.	510 TB patients	Residential area, gender, sputum smear, type of TB, HIV test, and weight loss.	Those did not start treatments or treatment was interrupted for two consecutive months or more.
Hill et al. (2005)	Prospective cohort, urban public clinics in Greater Banjul, Gambia.	301 newly diagnosed TB patients	Age, gender, ethnicity, type of TB, employment status, family history, knowledge of the disease, perceived benefit, pre-treatment cost, travel distance, and financial stress.	Failure to present for DOTS for three consecutive days.
Jenkins et al. (2013)	Retrospective cohort, Moldova	4021 TB patients	Residential area, homeless, gender, nationality, occupation, salary, educational level, history of detention, household size, number of children, staying with TB patient, the degree of lung pathology, sputum smear, sputum culture, HIV status, drug resistance, and region of residency.	Treatment interruption occurred for two consecutive months and more.
Jepchumba, et al. (2017)	Prospective cohort, 25 public centres (private and public), Nairobi County Kenya	291 newly diagnosed, PTB smear-positive patients	Gender, age, the source of income, employment status, the frequency of wage payment, Educational level, alcohol consumption, DOTS staff sufficiency, and nature of the facility	TB patients who did not start treatments or treatments were interrupted for two consecutive months or more.
Masini et al. (2016)	Retrospective cohort, Kenya	90,170 TB patients	Patient type, TB type, gender, age, body mass index (BMI), HIV status, type of DOTS, employment sector, and nutritional support.	TB patients who did not start treatments after diagnosis, or treatment interruption for two consecutive months.
Pefura-Yone et al. (2014)	Retrospective cohort, Yaounde Jamot Hospital, Cameroon	1688 TB patients	Age, gender, residence, place of screening, the setting of intensive phase treatment, Type of TB, Type of patient, HIV serology, and sputum smear conversion.	Treatment interruption occurred for two consecutive months and more.
Rutherford et al. (2013)	Prospective cohort, community lung clinic, Indonesia	249 TB patients	Household head and household demographics, patient health, treatment-seeking behaviour, clinic accessibility, TB knowledge, social support, previous experiences with TB, perceived stigma and experience with clinic staff and facilities.	Primary default was defined as treatment discontinuation for more than 2 weeks. Permanent default was treatment interruption for two consecutive months and more.

Table 2 *(continue)*

Author/Year	Study Design /Settings/ Study location	Participants	Exposure Assessment	Operational Definition of Treatment Interruption
Shargie & Lindtj (2007)	Prospective cohort, Southern Ethiopia (hospital and health centre)	404 newly diagnosed PTB smear-positive patients	Gender, age, marital status, education, occupation, family size, monthly family income, duration of symptoms, residential area, the zone of residence, distance to diagnostic centre, and distance to a treatment centre.	PTB patients on treatments for at least four weeks and treatments were interrupted for more than eight consecutive weeks.
Sylvère (2015)	Prospective cohort, Benin, West Africa	1226 TB patients	Age, gender, nationality, sputum smear, TB history, a form of TB, HIV/AIDS infection, TB regime, and DOTS territory	Two consecutive months of treatment discontinuation.

Table 3

Systematic review of timing and prognostic factors of TB treatment interruption

Author/ Year	Statistical Method	Outcome		Prognostic Factors Identified		
		TB Treatment Interruption Rate	Survival Time and Survival Probability	Factors	Hazard Ratio	95% Confidence Interval
Akessa et al. (2015)	Cox proportional hazard regression analysis	13.5%	Median survival time was 5.7 months. The survival probability at the end of treatment was 85.5%.	Residential area (rural address)	4.393	1.58-12.18
Hill et al. (2005)	Extended Cox Proportional hazard analysis	25.2%	The median time of treatment interruption was 115 days (IQR=21-195).	<u>0 to 90 days</u>		
				Perceived benefit (poor)	3.64	1.42–9.31
				<u>After 90 days of treatment</u>		
				Travel distance	2.67	1.05–6.81
Jenkins et al. (2013)	Cox proportional hazard Regression	14.7%	The median time of default was 110 days for newly diagnosed TB patients.	<u>Newly diagnosed cases</u>		
				Homeless	2.33	1.64- 3.29
				Increase the educational level	0.77	0.66- 0.91
				Living alone	1.57	1.20- 2.04
				Destructive lung pathology	1.59	1.25-2.02
				HIV positive	1.55	1.17- 2.05
				Presence of drug resistance	1.27	1.10- 1.46
				Previously TB treated patients	1.68	1.18-2.39
				Living together with TB patient	1.54	1.19-1.99

Table 3 *(continue)*

Author/ Year	Statistical Method	Outcome		Prognostic Factors Identified		
		TB Treatment Interruption Rate	Survival Time and Survival Probability	Factors	Hazard Ratio	95% Confidence Interval
Jepchumba et al. (2017)	Cox proportional hazard regression	6.5%	The median time of default was 56 days (IQR =36-105).	Type of health centre (public versus private)	0.210	0.05-0.95
				Adequate vs. Inadequate HCWs	0.195	0.07-.0.56
				Educational level (primary versus secondary)	5.28	1.18-23.59
Masini et al. (2016)	Mixed-effect cox proportional hazard modelling	4.5% for new patients	The hazard of treatment interruption was highest during the intensive phase.	Relapse cases	1.70	1.44–2.00
				Retreatment after failure	4.79	3.99–5.75
				Male	1.46	1.35–1.58
				Underweight versus normal	1.11	1.03–1.20
				Positive not on ART versus negative	1.96	1.70–2.26
Pefura-Yone et al. (2014)	Cox proportional hazard regression	20%	The median duration of treatment discontinuation was 90 days (IQR= 30-150).	Hospitalisation during the intensive phase	0.69	0.54-0.89
				Non-consenting for HIV screening	1.65	1.24-2.21
Rutherford, et al. (2013)	Cox proportional hazard regression	16%	The median time of default was 36 days (IQR= 7–99).	Liver disease	3.40	1.02-11.78
				Chest pain	2.25	1.06-4.77
				Night sweat	1.98	1.03-3.79
				Household wealth (poor)	4.24	1.12-16.09
				Walking to the clinic	4.53	1.39-14.71
				Low level of satisfaction with the clinic	3.85	1.17-12.62
Shargie & Lindtj (2007)	Cox proportional hazard regression	20%	The cumulative probability of default at the end of two months was 8.6%.	Distance from home to the treatment centre (>2 hours)	2.97	1.91-4.62
				Age (>25years)	1.71	1.09-2.68
Sylvère (2015)	Cox proportional hazard regression	5% for overall and 3% for intensive phase treatment interruption.	Median survival time was 2 months (IQR=1-3). The survival probability was 97.5% at 2 months and 94.9% at 6 months of treatment	Age	3.49	1.25-9.74
				HIV positive	3.82	2.28-6.41
				History of TB infection	0.13	0.03-0.53

interruption was 36 days (IQR=7–99) (Rutherford et al., 2013). Consistently, Masini et al. (2016) via a large retrospective cohort study involving 91099 TB patients in Kenya, had elicited evidence that the hazard rate of TB treatment interruption was highest during the intensive phase. Incongruent with preceding findings, Jepchumba et al. (2017) and Sylvère (2015) during their respective assessments at Kenya and West Africa, revealed that the median time of TB treatment interruption was towards the completion of the intensive phase.

Prognostic Factors In Terms of Time for TB Treatment Interruption

Prognostic factors in terms of time for TB treatment interruption elicited from the prognostic studies are presented in Table 3. These factors can be conveniently divided into socio-demographic and high-risk behaviours, clinical factors, health service factors as well as behavioural and cognitive factors.

Socio-demographic Factors and High-Risk Behaviour. In consistent with findings of Shargie and Lindtj (2007), and Sylvere (2015) demonstrated the prospective association between age and different time points of TB treatment interruption even though both had contradicting findings. Meanwhile, several other studies demonstrated that the age factor is an important confounder to be addressed during TB treatment interruption. Similarly, gender has also exerted confounding effect in most of the studies. Despite extensive evaluation of the demographic and socioeconomic profiling in most of the studies, there has been inconsistent findings across residential area, educational level, and family income. In the lens of high-risk behaviours, there is a limited evaluation of the smoking status, illegal drug use as well as alcohol consumption.

Clinical Characteristics. In this review, clinical parameters such as underlying disease, treatment-related factors, and comorbidity information were widely studied particularly in retrospective cohort studies. In essence, there have been consistent influences demonstrated by the type of TB patients and HIV infection on the time of TB treatment interruption, in which those with previous history of TB treatment and HIV positive status show higher probability of defaulting (Jenkins et al., 2013; Masini et al., 2016; Sylvere, 2015). In addition, there have been inconsistent findings in term of the history of hospitalization, body weight, symptoms improvement, sputum culture, chest x-ray grading as well as comorbidities.

Cognitive and Behavioural Factors. Despite limited assessment of cognitive and behavioural factors influence on the time of TB treatment interruption, Hill et al. (2005) evidently demonstrated that among 301 PTB patients in Gambia, those who were uncertain about the effectiveness of the standard treatment were 3.64 times at risk of developing treatment interruption as compared to the reference group. However, this effect varied across the time, in which perceived benefits was the only exerting and significant effect on the time of TB treatment interruption in the first 90 days of the treatment course, but not thereafter. On the other hand, Rutherford et al. (2013) attempted a temporal assessment of TB knowledge and perceived stigma, but there were no significant influences to TB treatment interruption observed in terms of TB treatment interruption.

Health Service Factors. Health service factors such as staff adequacy, perceived access to healthcare centre and perceived patients' satisfaction have gained much interest from researchers during the assessment of the time of TB treatment interruption. It was evident in three distinct studies that health service access was found to have a significant association with different times of treatment interruption (Hill et al., 2005; Rutherford et al., 2013; Shargie & Lindtj, 2007). Earlier, Hill et al. (2005) in a prospective cohort study involving 301 TB patients in an urban community in Gambia, reported that patients with a travel time of more than half an hour or spent more than six *dalasis* (0.16 euro) for travel access had a higher risk of default, which was significantly observed after three months of treatment. In this study, the author postulated the correlation of built-up cost over time with a higher rate of defaulting during the maintenance phase. In contrast, a temporal assessment among 249 TB patients in Indonesia showed that travel time, travel distance (in kilometres) and travel cost did not influence TB treatment interruption significantly (Rutherford et al. (2013). Instead, patients who walked to the treatment centre had a higher risk of early treatment interruption, as the median time of default was 36 days. In the meantime, Shargie and Lindtj (2007) depicted that walking time of more than two hours posed a risk of maintenance phase treatment interruption in the rural community of Southern Ethiopia. The two latter studies shared a similarity, as both evaluated rural communities who predominantly walked to the treatment centres.

DISCUSSION

In this review, most of the studies shared several common characteristics, in which the selected studies were conducted in developing and high TB burden countries, and mean age of recruited participants was within the middle age group. However, quantitative analysis could not be performed due to limited finalized studies and considerable diversity across the studies, particularly on the study designs, duration of follow-up, as well as the operational definition of TB treatment interruption. Two studies did not explicitly report on the timing of default. Instead, the cumulative hazard risk or hazard functions were presented (Marsini et al., 2016; Shargie & Lindtj, 2007).

Notwithstanding the foregoing restriction, this review revealed that the timing of treatment interruption predominantly occurred during the maintenance phase. This finding is incongruent with a previous systematic review conducted across the temporal data from developing countries which also demonstrated the maintenance phase as the point of exit from the treatment course (Kruk et al., 2008). In addition, this has been supported by previous qualitative studies whereby it was highlighted that most of the patients had significant symptoms improvement by the time they were in the maintenance phase (Martins et al., 2008; Widjanarko et al., 2009). Having said that, Rutherford et al. (2013) demonstrated that the median time of the TB treatment interruption was 36 days (IQR=7-99), whilst Masini et al. (2016) pointed out that the hazard rate of TB treatment interruption was highest during intensive phase, which collectively raised a postulation that additional effort for case holding relies on rigorous policy implementation and effectiveness of the existing programme in individual country to ensure adherence.

In the context of prognostic factors of TB treatment interruption, various factors contributing to TB treatment interruption at different time points have been identified. These factors are framed via socio-demographic and high-risk behaviour, clinical characteristics, health service factors as well as cognitive and behavioural factors. Prominently, travel distance and HIV infection were demonstrated as one of the prognostic factors of time to TB treatment interruption (Hill et al., 2005; Jenkins et al., 2013; Masini et al., 2016; Rutherford et al., 2013; Shargie & Lindtj, 2007; Sylvere, 2015). Undoubtedly, there are inconsistent findings, which largely influenced by study settings and study participants. Several studies included those with comorbidities such as HIV infection, liver disease, and diabetes which were predisposed to the complexity of treatment responses and standard care, thus influenced the mean time of duration of follow up. The comorbidities factors could also contribute to the competing risk of event occurrence, whereby this would lead to overestimation of the hazard of other factors.

Above all, findings from this review should be interpreted with caution. Firstly, this review was confined to published studies in which the generalizability of study population may render some restrictions. Crucially, the large heterogeneity was observed across the published studies particularly on study design, study participants and study population, duration of follow up, sample size and operational definition of TB treatment interruption. For example, there was considerable variation of TB treatment interruption rate, which strongly depending on the duration of follow up to allow sufficient event to be encountered, as well as the diversity of baseline characteristics of recruited participants. Therefore, a meta-analysis could not be performed to estimate the aggregate effect measure. In addition, this review included two studies dated more than 10 years. In this light, existing national programs and treatment protocol may have evolved and advances thereby impaired comparability of the studies.

Fundamentally, this review sheds light on the gap of individual research on longitudinal assessment of TB treatment interruption. It was demonstrated that most researchers focus on demographic and biomedical context influences. While behavioural and psychosocial contexts have been echoed by WHO as a crucial aspect to be explored in TB treatment interruption, the prospective assessment of TB treatment interruption at different time points of the treatment course should emphasize on high risk behaviour, cognitive and behavioural aspect as well as psychosocial context which has been widely demonstrated in the literature as the determinants of TB treatment interruption (Sabate, 2003; World Health Organization, 1999). In addition, the assessment of events during the course of treatment, such as drug side effects and care provider-patient interactions, which could considerably influence a patient's decision to continue treatment, was not demonstrated in this review. This gap finding should be further explored, hence to endeavour a holistic assessment of health services factors which could influence TB treatment interruption. In terms of analysis, the quality of statistical reporting in future prognostic studies can be enhanced by subgroup analysis and by allowing a competing risks analysis of the hazard estimation in catering for diverse participants and organizational characteristics.

CONCLUSION

Although this review generally reveals that TB treatment interruption occurs more frequently during the maintenance phase, there has been a concomitant gap in understanding the prognostic factors of different time points of TB treatment interruption. The heterogeneity across the studies particularly on study design and operational definition of TB treatment interruption may limit solid inferences on the outcome measure, which requires further evaluation. In essence, health practitioners should integrate time-related information, while furthering suggestion that prognostic studies should incorporate assessments of cognitive, behavioural studies and psychosocial evaluation. The findings would be able to provide further evidence in framing future public health strategy.

ACKNOWLEDGMENT

The authors would like to thank the Director General of Health Malaysia for permission to publish this manuscript.

REFERENCES

Akessa, G. M., Tadesse, M., & Abebe, G. (2015). Survival analysis of loss to follow-up treatment among tuberculosis patients at Jimma University Specialized Hospital, Jimma, Southwest Ethiopia. *International Journal of Statistical Mechanics, 2015*, 1-7.

Bagchi, S., Ambe, G., & Sathiakumar, N. (2010). Determinants of poor adherence to anti-tuberculosis treatment in Mumbai, India. *International Journal of Preventive Medicine, 1*(4), 223-232.

Cerdá, J. M. V., Conde, M. T. M., Verdugo, R. M., Yébenes, M., & Gómez, M. A. C. (2014). Adherence, satisfaction and health-related quality of life in HIV-infected patients with antiretroviral treatment in Spain: ARPAS study. *Hospital Pharmacy, 38*(4), 291-299.

Chandrasekaran, V., Gopi, P. G., Subramani, R., Thomas, A., Jaggarajamma, K., & Narayanan, P. R. (2005). Default during the intensive phase of treatment under DOTS programme. *Indian Journal of Tuberculosis*, *52*(44), 197-202.

Chee, C. B., Boudville, I. C., Chan, S. P., Zee, Y. K., & Wang, Y. T. (2000). Patient and disease characteristics, and outcome of treatment defaulters from the Singapore TB control unit--a one-year retrospective survey. *The International Journal of Tuberculosis and Lung Disease : The Official Journal of the International Union against Tuberculosis and Lung Disease*, *4*(6), 496-503.

Connolly, C., Davies, G. R., & Wilkinson, D. (1999). Who fails to complete tuberculosis treatment? Temporal trends and risk factors for treatment interruption in a community-based directly observed therapy programme in a rural district of South Africa. *International Journal of Tuberculosis and Lung Disease, 3*(12), 1081-1087.

Dodor, E. A. (2004). Tuberculosis treatment default at the Communicable Diseases Unit of Effia-Nkwanta Regional Hospital: A 2-year experience. *International Journal of Tuberculosis and Lung Disease, 8*(11), 1337-1341.

Dominguez-Castellano, A., Muniain, M. A., Rodriguez-Bano, J., Garcia, M., Rios, M. J., Galvez, J., & Perez-Cano, R. (2003). Factors associated with time to sputum smear conversion in active pulmonary tuberculosis. *The International Journal of Tuberculosis and Lung Disease, 7*(5), 432-438.

da Silva Garrido, M., Penna, M. L., Perez-Porcuna, T. M., de Souza, A. B., da Silva Marreiro, L., Albuquerque, B. C., ... & Bührer-Sékula, S. (2012). Factors associated with tuberculosis treatment default in an endemic area of the Brazilian Amazon: A case control-study. *PLoS ONE, 7*(6), 1-7.

Hayden, J. A., Cote, P., Steenstra, I. A., & Bombardier, C. (2008). Identifying phases of investigation helps planning, appraising, and applying the results of explanatory prognosis studies. *Journal of Clinical Epidemiology, 61*(6), 552-560.

Hill, P. C., Stevens, W., Hill, S., Bah, J., Donkor, S. A., Jallow, A., & Lienhardt, C. (2005). Risk factors for defaulting from tuberculosis treatment: A prospective cohort study of 301 cases in the Gambia. *International Journal of Tuberculosis and Lung Disease, 9*(12), 1349-1354.

Jenkins, H. E., Ciobanu, A., Plesca, V., Crudu, V., Galusca, I., Soltan, V., & Cohen, T. (2013). Risk factors and timing of default from treatment for non-multidrug-resistant tuberculosis in Moldova. *The International Journal of Tuberculosis and Lung Disease, 17*(3), 373-380.

Jepchumba, V., Karanja, S., Amukoye, E., Muthami, L., & Kipruto, H. (2017). Timing and determinants of tuberculosis treatment interruption in Nairobi County, Kenya. International *Journal of Public Health Science (IJPHS), 6*(3), 203-212.

Kleinbaum, D. G. (1996). *Survival analysis: a self-learning text (statistics in the health sciences)*. New York, NY: Springer-Verlag.

Kliiman, K., & Altraja, A. (2010). Predictors and mortality associated with treatment default in pulmonary tuberculosis. *International Journal of Tuberculosis and Lung Disease, 14*(4), 454-463.

Kruk, M. E., Schwalbe, N. R., & Aguiar, C. A. (2008). Timing of default from tuberculosis treatment: a systematic review. *Tropical Medicine and International Health, 13*(5), 703-712.

Kuaban, C., Bame, R., Mouangue, L., Djella, S., & Yomgni, C. (2009). Non conversion of sputum smears in new smear positive pulmonary tuberculosis patients in Yaoundé, Cameroon. *East African Medical Journal, 86*(5), 219-225.

Maher, D. (2009). The natural history of mycobacterium tuberculosis infection in adults. In H. S. Schaaf & A. Zumla (Eds.), *Tuberculosis: A Comprehensive Clinical Reference* (pp. 129-132). London, UK: Elsevier.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Journal of Clinical Epidemiology, 62*(10), 1006-1012.

Martins, N., Grace, J., & Kelly, P. M. (2008). An ethnographic study of barriers to and enabling factors for tuberculosis treatment adherence in Timor Leste. *International Journal of Tuberculosis and Lung Disease, 12*(5), 532-537.

Marx, F. M., Dunbar, R., Enarson, D. A., & Beyers, N. (2012). The rate of sputum smear-positive tuberculosis after treatment default in a high-burden setting: a retrospective cohort study. *PLoS ONE, 7*(9), 1-9.

Masini, E. O., Mansour, O., Speer, C. E., Addona, V., Hanson, C. L., Sitienei, J. K., … & Mungai, B. N. (2016). Using survival analysis to identify risk factors for treatment interruption among new and retreatment tuberculosis patients in Kenya. *PLoS ONE, 11*(10), 1-19.

Millet, J. P., Orcau, A., de Olalla, P. G., Casals, M., Rius, C., & Cayla, J. A. (2009). Tuberculosis recurrence and its associated risk factors among successfully treated patients. *Journal of Epidemiology and Community Health, 63*(10), 799-804.

Nahid, P., Jarlsberg, L. G., Rudoy, I., de Jong, B. C., Unger, A., Kawamura, L. M., ... & Daley, C. L. (2011). Factors associated with mortality in patients with drug-susceptible pulmonary tuberculosis. *BMC Infectious Diseases, 11*(1), 1-7.

Naing, N. N., D'Este, C., Isa, A. R., Salleh, R., Bakar, N., & Mahmod, M. R. (2001). Factors contributing to poor compliance with anti-TB treatment among tuberculosis patients. *Southeast Asian Journal of Tropical Medicine and Public Health, 32*(2), 369-382.

Pablos-Méndez, A., Sterling, T. R., & Frieden, T. R. (1996). The relationship between delayed or incomplete treatment and all-cause mortality in patients with tuberculosis. *JAMA : The Journal of the American Medical Association, 276*(15), 1223-1228.

Pefura-Yone, E. W., Kengne, A. P., & Kuaban, C. (2014). Non-conversion of sputum culture among patients with smear positive pulmonary tuberculosis in Cameroon: a prospective cohort study. *BMC infectious Diseases, 14*(1), 1-6.

Pettit, A. C., Cummins, J., Kaltenbach, L. A., Sterling, T. R., & Warkentin, J. V. (2013). Non-adherence and drug-related interruptions are risk factors for delays in completion of treatment for tuberculosis. *International Journal of Tuberculosis and Lung Disease, 17*(4), 486-492.

Rutherford, M. E., Hill, P. C., Maharani, W., Sampurno, H., & Ruslami, R. (2013). Risk factors for treatment default among adult tuberculosis patients in Indonesia. *The International Journal of Tuberculosis and Lung Disease, 17*(10), 1304-1309.

Sabate. (2003). *Adherence to long-term therapies: evidence for action*. Geneva: World Health Organization. Retrieved June 9, 2018, from http://apps.who.int/iris/bitstream/10665/42682/1/9241545992.pdf

Shargie, E. B., & Lindtjørn, B. (2007). Determinants of treatment adherence among smear-positive pulmonary tuberculosis patients in Southern Ethiopia. *PLoS Medicine, 4*(2), 0280-0287.

Sulaiman, S. A. S., & Ali, J. A. (2010). Clinical modalities and therapeutic outcomes; observational cohort comparison between ever-smokers versus never-smokers of tuberculosis patients in Penang, Malaysia. *European Journal of General Medicine, 7*(4), 389-397.

Sylvère, T. A. (2015). Default time from tuberculosis treatment in the southern republic of benin using mixture cure model for survival analysis. *Biometrics and Biostatistics International Journal, 2*(5), 1-13.

Tachfouti, N., Slama, K., Berraho, M., Elfakir, S., Benjelloun, M. C., El Rhazi, K., & Nejjari, C. (2013). Determinants of tuberculosis treatment default in Morocco: results from a national cohort study. *The Pan African Medical Journal, 14*, 1-7.

Vijay, S., Kumar, P., Chauhan, L. S., Vollepore, B. H., Kizhakkethil, U. P., & Rao, S. G. (2010). Risk factors associated with default among new smear positive TB patients treated under DOTS in India. *PLoS ONE, 5*(4), 1-9.

Vree, M., Huong, N. T., Duong, B. D., Sy, D. N., Van, L. N., Co, N. V, … & Borgdorff, M. W. (2007). Mortality and failure among tuberculosis patients who did not complete treatment in Vietnam: a cohort study. *BMC Public Health, 7*(1), 1-8.

Widjanarko, B., Gompelman, M., Dijkers, M., & van der Werf, M. J. (2009). Factors that influence treatment adherence of tuberculosis patients living in Java, Indonesia. *Patient Preference and Adherence, 3*, 231-238.

World Health Organization. (2017). *Global Tuberculosis Report 2017*. Retrieved February 12, 2018, from http://apps.who.int/iris/bitstream/10665/259366/1/9789241565516-eng.pdf?ua=1

World Health Organization. (2014). *Definitions and reporting framework for tuberculosis – 2013 revision*. Geneva, Switzerland: World Health Organization.

World Health Organization. (1999). *Fixed-dose combination tablets for the treatment of tuberculosis*. Geneva. Retrieved May 4, 2018, from http://apps.who.int/iris/bitstream/10665/65981/1/WHO_CDS_CPC_TB_99.267.pdf

Pertanika J. Sci. & Technol. 27 (4): 1527 - 1538 (2019)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Simulation Work on Blood Glucose Control for Type 1 Diabetes using Modified Hovorka Equations

Ayub Md Som[1,2]*, Nur Farhana Mohd Yusof[1], Sherif Abdulbari Ali[1] and Nurafiqah Zawawi[1]

[1]Faculty of Chemical Engineering, Universiti Teknologi MARA, 40450 UiTM, Shah Alam, Selangor, Malaysia
[2]Accounting Research Institute, Universiti Teknologi MARA, 40450 UiTM, Shah Alam, Selangor, Malaysia

ABSTRACT

Failure of pancreas can cause uncontrolled blood glucose levels in the body. This research focuses on type 1 diabetes patients who depend on external insulin injection. The Hovorka model was used as the mathematical model in the development of control algorithm for artificial pancreas. However, the model showed a lack of interaction on selected parameters and variables in its glucose-insulin dynamic system. An improvement on the Hovorka equations was done, but no work was carried out to simulate the proposed equations. The objectives of this study are to simulate the modified Hovorka equations using MATLAB and to compare the simulation results between the reference and modified ones. This study showed better interaction among all variables and parameters on its glucose-insulin dynamic system using the modified equations compared to the original equations. The lower administered amount of insulin, U_t at 16.7mU/min and 20mU/min could regulate the blood glucose level at normoglycemic condition throughout the study.

Keywords: Artificial pancreas, control algorithm, hovorka model, MATLAB simulation, type 1 diabetes

ARTICLE INFO

Article history:
Received: 9 April 2018
Accepted: 13 May 2019
Published: 21 October 2019

E-mail addresses:
ayub522@salam.uitm.edu.my, ayub522@uitm.edu.my (Ayub Md Som)
nurfarhana2002@yahoo.com (Nur Farhana Mohd Yusof)
abdulbari@salam.uitm.edu.my (Sherif Abdulbari Ali)
nurafiqahzawawi93@gmail.com (Nurafiqah Zawawi)
*Corresponding author

INTRODUCTION

Artificial pancreas or closed-loop system consists of continuous subcutaneous insulin infusion (CSII) pump, continuous glucose monitoring (CGM) sensor, CGM receiver and control algorithm which measure and regulate current blood glucose level of type

ISSN: 0128-7680
e-ISSN: 2231-8526

1 diabetes patients in an automated manner so as to minimise hyperglycaemia without increasing hypoglycaemia (Elleri et al., 2011; Forlenza et al., 2016; Messori et al., 2015; Thabit & Hovorka, 2012; Thabit & Hovorka, 2016). As control algorithm is one of the most important components which act as a heart of the system, the controller should be able to communicate with all components in the device to control insulin infusion activity and blood glucose level in the patient's body (Hovorka et al., 2006; Hovorka et al., 2010).

The device is needed to function as a fully automated device in order to reduce the dependency on manual insulin injection by the patients. The mathematical equations of Hovorka model (Hovorka et al., 2004) are used in the control algorithm; however, the model somewhat lacks interaction and interrelation of selected parameters in its glucose-insulin dynamics, specifically involving its glucose subsystem, plasma insulin subsystem, and insulin action subsystem (Yusof et al., 2012; Yusof et al., 2013; Yusof et al., 2014; Yusof et al., 2015). In the Hovorka model as shown in Figure 1, it seems that only insulin on action transport (x_1) and insulin on endogenous production (x_3) interact in mass of glucose in the accessible compartment (Q_1). Meanwhile the mass of glucose in the non-accessible compartment (Q_2), interaction only occurs with insulin on action transport (x_1) and insulin on action disposal (x_2). As such, the modified Hovorka equations have been proposed by Yusof et al. (2012) in order to improve the interaction and interrelation of all related parameters in the glucose-insulin dynamics system. However, at present no work has been carried out to simulate the proposed equations in order to obtain its performance and user-ability. Thus, this study attempts to show the simulation work by using the improved equations.

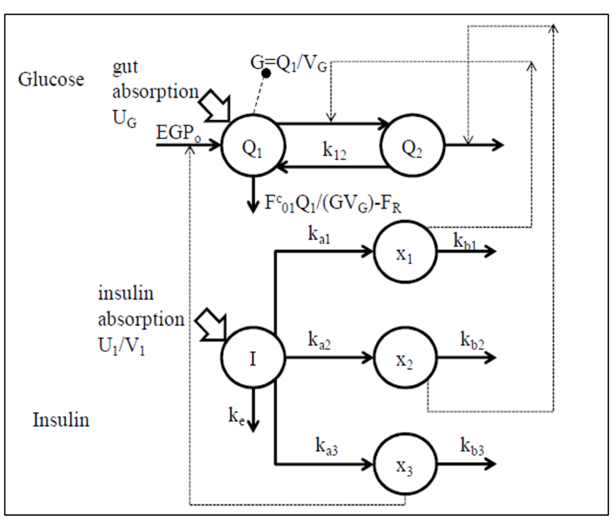


Figure 1. Schematic diagram of Hovorka model

METHODOLOGY

Hovorka Model with Modified Equations

Some of the mathematical equations from the Hovorka model (Hovorka et al., 2004) had been changed and transformed into a set of equations. The transformation of modified Hovorka equations using system identification techniques as previously stated by Yusof et al. (2014) can be clearly seen in Figure 2. As shown in the figure, it depicts a solid improvement in interactions and interrelations of all parameters and variables involved in glucose-insulin dynamics system using the modified Hovorka equations. By employing the system identification technique and modifying related equations, all insulin action subsystems completely interact with the equations of mass of glucose in the accessible compartment and mass of glucose in the non-accessible compartment. Insulin on action transport (x_1), insulin on action disposal (x_2) and insulin on endogenous production (x_3) can reach the mass of glucose in the accessible compartment (Q_1), and non-accessible compartment (Q_2) interactions.

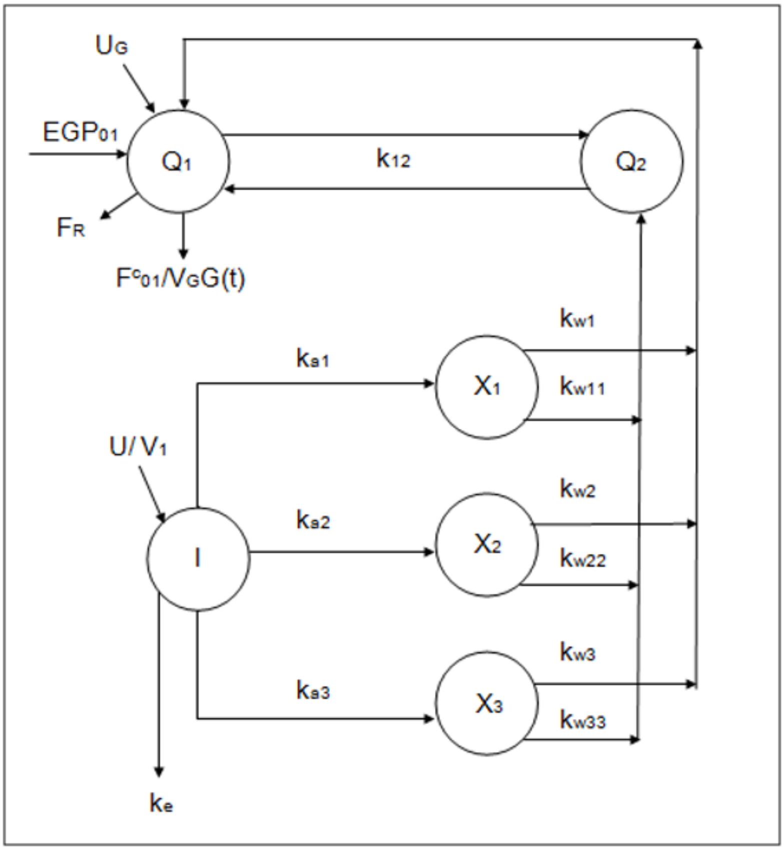


Figure 2. Schematic diagram of Hovorka model with modified equations

Subsequently, some related equations in the Hovorka model have been modified in certain subsystems of the model. The modified equations take place, namely in glucose subsystem, plasma insulin concentration, and insulin action subsystem. Meanwhile, the other equations remain unchanged.

Glucose subsystem has been improved by adding all insulin action variables in the equations for accessible compartment and non-accessible compartment as shown in equations (1) and (2) as follows:

$$\frac{dQ_1}{dt} = EGP_o + U_G + 0.01\,Q_2 + [x_1 k_{w1} + x_2 k_{w2} + x_3 k_{w3}] - F_R\,Q_1 - \left[\frac{F_{01}^c}{V_G G(t)}\right] Q_1 - 0.002 Q_1 \quad [1]$$

$$\frac{dQ_2}{dt} = [k_{w11} x_1(t) + k_{w22} x_2(t) + k_{w33} x_3(t)] + EGP_o[k_{w1} x_1(t) + k_{w2} x_2(t) + k_{w3} x_3(t)] k_{12}\,Q_2 \quad [2]$$

Q_1 and Q_2 represent the glucose mass in the accessible and non-accessible compartments, respectively. The constants of k_{w1}, k_{w2}, k_{w3}, k_{w11}, k_{w22} and k_{w33} represent the transfer rate constants of insulin action subsystem. Meanwhile, the constant of k_{12} is represented as transfer rate from non-accessible to accessible compartment. EGP_0 represents endogenous glucose production (EGP) that is extrapolated to the zero insulin concentration. U_G is represented as the absorption amount of glucose into the blood vessel. The parameter of F_c^{01} is the total of non-insulin dependent glucose flux and F_R is represented as the renal glucose clearance (Hovorka et al., 2004).

$$F_R = \begin{cases} 0.003\,(G-9)V_G if G \geq 9\ mmolL^{-1} \\ 0 \quad otherwise \end{cases}$$

$$F_C^{01} = \begin{cases} F_{01} if G \geq 4.5\ mmolL^{-1} \\ \frac{F_{01} G}{4.5} otherwise \end{cases}$$

The equations in the insulin subsystem remain the same as the Hovorka model. Equations (3) and (4) show the insulin subsystem in the accessible and non-accessible compartments. S_1 and S_2 represent as the insulin sensitivity in the accessible and non-accessible compartments, respectively.

$$\frac{ds_1(t)}{dt} = u(t) - \frac{s_1(t)}{t_{max,I}} \quad [3]$$

$$\frac{ds_2(t)}{dt} = \frac{s_1(t)}{t_{max,I}} - \frac{s_2(t)}{t_{max,I}} \quad [4]$$

The variables of insulin action are also added in plasma insulin concentration equation. The plasma insulin concentration I(t) is changed as shown in equation (5) in which k_e is the fractional elimination rate, V_I is the distribution volume and U_I is represented as the production amount of insulin required into the blood vessel.

$$\frac{dI(t)}{dt} = \left[\frac{U_I(t)}{V_I}\right] - k_e I(t) - [k_{w1} x_1(t) + k_{w2} x_2(t) + k_{w3} x_3(t)] \quad [5]$$

Meanwhile, equations (6) to (8) represent insulin action subsystem equations on action transport, action disposal and endogenous production, respectively. The constant of k_{a1}, k_{a2},$k_{a3,}$ k_{w1}, $k_{w11,}$ k_{w2}, $k_{w22,\,,}$ k_{w3}, and k_{w33} are the deactivation and activation rates of insulin action.

$$\frac{dx_1}{dt} = -k_{a1}x_1(t) + k_{w1}I(t) + k_{w11}I(t) \qquad [6]$$

$$\frac{dx_2}{dt} = -k_{a2}x_2(t) + k_{w2}I(t) + k_{w22}I(t) \qquad [7]$$

$$\frac{dx_3}{dt} = -k_{a3}x_3(t) + k_{w3}I(t) + k_{w33}I(t) \qquad [8]$$

The constants and parameters are included in the equations to represent the glucose absorption for type 1 diabetes conditions. The constants and parameters in the equations are determined by its specific values. The constant values and parameters are defined as shown in Tables 1 and 2, respectively (Yusof et al., 2012; Yusof et al., 2013; Yusof et al., 2014).

Table 1

The constant values used for modified hovorka equations

Symbol	Constant	Value & Unit
k_{12}	Transfer rate	0.066 min^{-1}
k_{a1}	Deactivation rate	0.006 min^{-1}
k_{a2}	Deactivation rate	0.06 min^{-1}
k_{a3}	Deactivation rate	0.03 min^{-1}
k_{w1}	Activation rate	50.1 min^{-1}
k_{w11}	Activation rate	-10 min^{-1}
k_{w2}	Activation rate	50.1 min^{-1}
k_{w22}	Activation rate	-0.01 min^{-1}
k_{w3}	Activation rate	50.1 min^{-1}
k_{w33}	Activation rate	-0.01 min^{-1}
k_e	Insulin elimination from plasma	0.138 min^{-1}
V_I	Insulin distribution volume	0.12 L kg^{-1}
V_G	Glucose distribution volume	0.16 L kg^{-1}
A_G	Carbohydrate(CHO) bioavailability	0.8 (unit less)
$t_{max,G}$	Time-to-maximum of CHO absorption	40 min

Table 2

The parameter values used for modified hovorka equations

Symbol	Parameter	Value & Unit
S^f_{IT}	Insulin sensitivity of distribution/ transport	51.2 x 10^{-4} min^{-1} per mU L^{-1}
S^f_{ID}	Insulin sensitivity of disposal	8.2 x 10^{-4} min^{-1} per mU L^{-1}

Table 2 *(Continued)*

Symbol	Parameter	Value & Unit
EPG_o	EGP extrapolated to zero insulin concentration	0.0161 $mmolkg^{-1}$ min^{-1}
F_{01}	Non-insulin-dependent glucose flux	0.0097 $mmolkg^{-1}$ min^{-1}
$t_{max,1}$	Time-to-maximum of absorption of subcutaneously injected short-acting insulin	55 min

RESULTS AND DISCUSSION

Simulation work has been done through MATLAB programming as discussed in this section.

Plasma Glucose Concentration

The plasma glucose concentration is simulated based on the amount of insulin administered, U_t. The trends of plasma glucose concentration simulated at U_t of 16.7 mU/min, 20 mU/min, 50 mU/min, 75 mU/min and 100 mU/min are shown in Figure 3.

As shown in Figure 3, the plasma glucose concentration is initialized at 4.5 mmol/L of plasma glucose. The concentration of plasma glucose increased rapidly to the peak point. At certain times, the glucose concentration reduced and became constant. The concentration of plasma glucose achieved normoglycemic condition (4 mmol/L – 6 mmol/L) when the U_t = 16.7 mU/min and U_t = 20 mU/min. On the contrary from the previous study as shown in Figure 4, the simulation of plasma glucose concentration by using Hovorka model with U_t = 16.7 mU/min showed that the concentration was too high and could cause hyperglycaemia (Yusof et al., 2015). The difference in both simulation results was due to the improved model from glucose subsystem in accessible compartment and non-accessible compartment as stated by Yusof et al. (2012).

Both U_t = 16.7 mU/min and U_t = 20 mU/min were regulated at constant blood glucose level at different times. When U_t = 16.7 mU/min was infused to the body, the blood glucose level of 5 mmol/L was achieved at 343 minutes and it remained constant throughout the simulation work. However, the infusion of insulin to the body at U_t = 20 mU/min showed that the blood glucose level of 4 mmol/L was achieved at 210 minutes and it remained constant thereafter. Thus, the 20 mU/min of insulin administered had achieved constant blood glucose level faster than the 16.7 mU/min of insulin. The plasma glucose concentrations for 50 mU/min, 75 mU/min and 100 mU/min of insulin administered had reached lower peak glucose concentrations compared to 16.7 mU/min and 20 mU/min. These amounts of administered insulin had reached constant blood glucose level at earlier times than the lower administered insulin. However, this condition could lead to prolonged hypoglycaemia due to constantly lower blood glucose levels (less than 4 mmol/L) throughout the simulation work (Bilous et al., 2010). From the previous Hovorka model simulation as shown in Figure 4, the higher administered insulin amount showed that the blood glucose level was regulated at normoglycemia level (Yusof et al., 2015). In this study, it shows that the lower administered insulin amount can regulate at normoglycemia level but it takes some time.

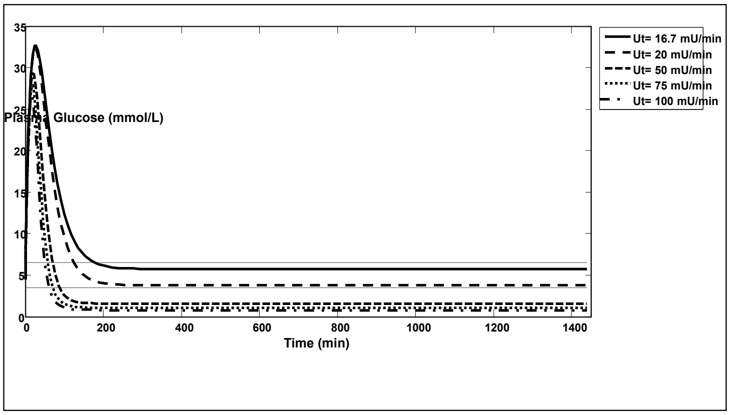


Figure 3. Plasma glucose concentration using modified Hovorka equations

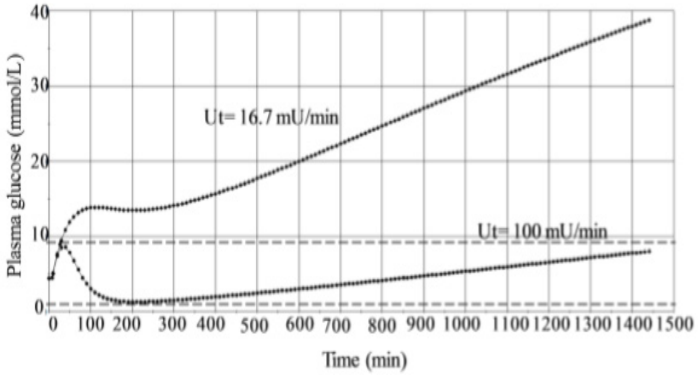


Figure 4. Plasma glucose concentration using Hovorka model

Plasma Insulin Concentration

The plasma insulin concentration was identified by simulating U_t at 16.7 mU/min, 20 mU/min, 50 mU/min, 75 mU/min and 100 mU/min, respectively. The trends of plasma insulin concentration simulated at these administered U_t are shown in Figure 5.

As shown in Figure 5, the plasma insulin concentration started at 15 mU/L of plasma insulin. The concentration of plasma insulin increased with time and it remained constant thereafter at certain times for all the administered insulins under study.

The plasma insulin concentration reached 900 mU/L and remained constant after 300 minutes for U_t at 16.7 mU/min. Meanwhile, the simulation work for plasma insulin concentration for U_t at100 mU/min achieved its steady state (remained constant) at 400 minutes and reached as high as 5500 mU/L of its plasma insulin concentration. Thus, the lower amount of insulin infused to the body would cause the plasma insulin concentration to be lower than when the higher amount of insulin was administered. In addition, the time taken for plasma insulin concentration to achieve its steady state was lower when the insulin administered amount was low.

Different administered insulin amounts have shown different performances as depicted in Figure 5. This was due to the different amount of insulin infused in the body. When the infusion rate of insulin was higher, the plasma insulin concentration became higher and the time taken for the plasma insulin concentration to reach its steady state was longer.

The trend shown was the same as previously studied by Yusof et al. (2015) through simulation works using the Hovorka equations as shown in Figure 6. Plasma insulin concentration increased as the time increased. However, the plasma insulin concentration for the modified Hovorka equations reached its steady state at a shorter time as compared to the one in the previous work by Yusof et al. (2015). This result was in agreement with the simulation work as carried out by Yusof et al. (2012) using the modified Hovorka equations for x_1, x_2 and x_3 as shown in Figures 7, 8 and 9, respectively.

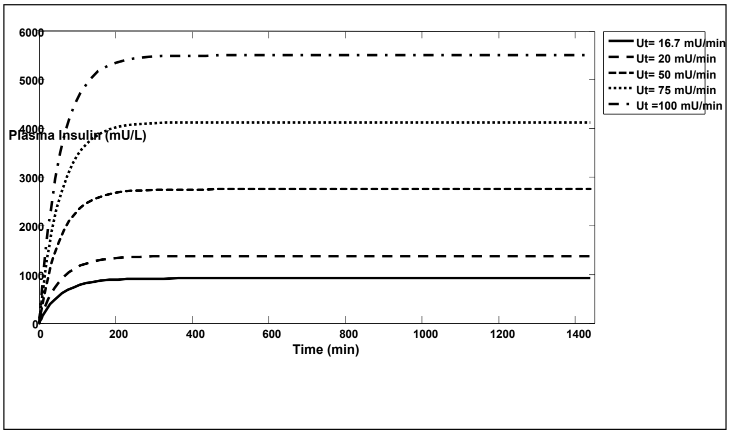


Figure 5. Plasma Insulin concentration using modified Hovorka equations

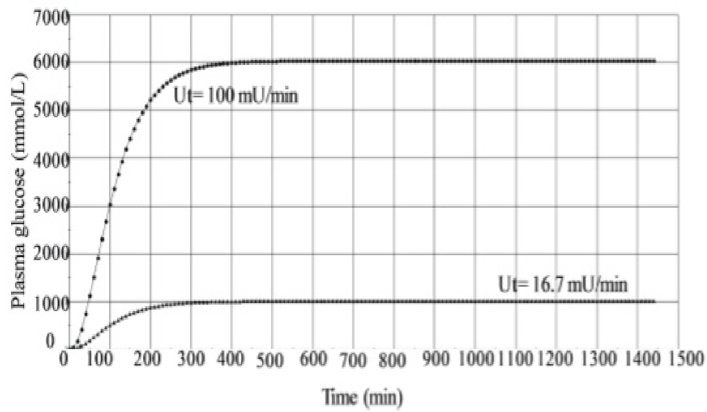
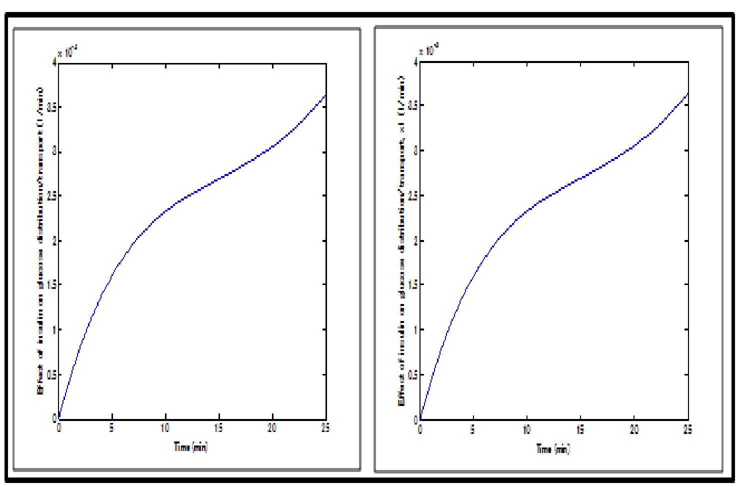
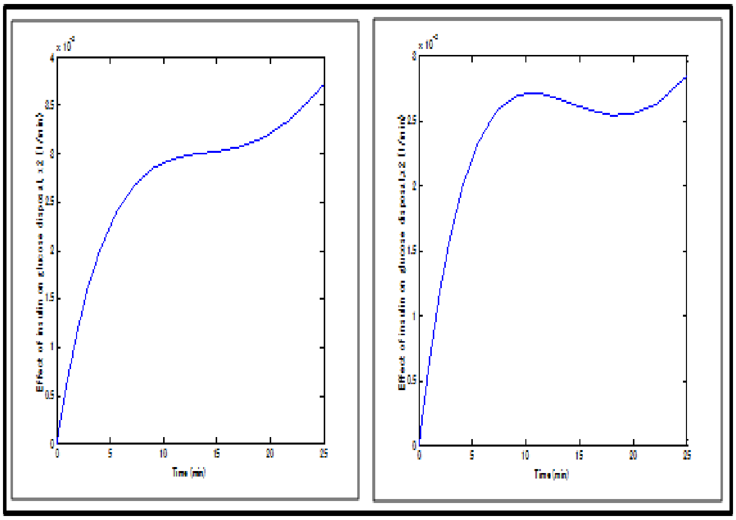


Figure 6. Plasma insulin concentration using Hovorka model equations



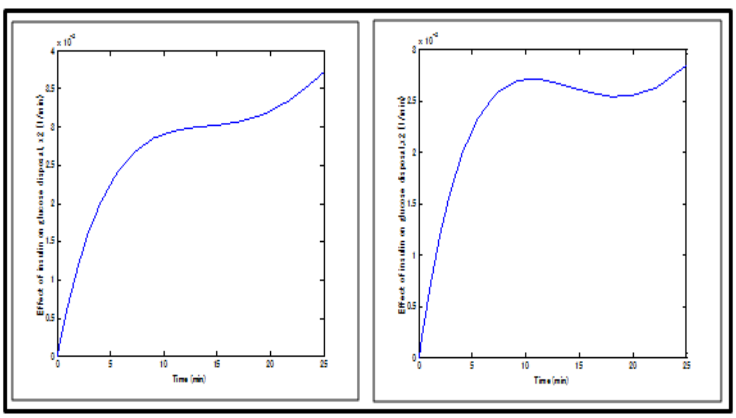
Hovorka model Modified Hovorka

Figure 7. Effect of insulin on glucose distribution/transport, x_1 between Hovorka model and modified Hovorka equations



Hovorka model Modified Hovorka

Figure 8. Effect of insulin on glucose disposal, x_2 between Hovorka model and modified Hovorka equations



Hovorka model Modified Hovorka equations

Figure 9. Effect of insulin on endogenous glucose production, x_3 between Hovorka model and modified Hovorka equations

CONCLUSION

The present study showed simulation work for the modified Hovorka equations so as to improve the interaction and interrelation among all parameters and variables on the glucose-insulin dynamic system. The U_t was the most active parameter. Thus, two insulins administered, U_t namely at 16.7 mU/min and 20 mU/min were chosen as they gave the best effect on glucose-insulin dynamic system.

The analysis showed that these lower U_t at 16.7 mU/min and 20 mU/min could regulate the blood glucose level at 5 mmol/L and 4 mmol/L, respectively. This condition was called normoglycemia in which the blood glucose level was in the safe range. Meanwhile, the higher U_t might have led to hypoglycaemia due to lower blood glucose level, although they had achieved a constant blood glucose level faster than the higher U_t. In addition, the plasma insulin concentration increased as time increased when the U_t was infused to the body. It achieved its steady state after 300 minutes and 400 minutes for U_t at 16.7 mU/min and 100 mU/min, respectively. The higher U_t infused to the body would cause the plasma insulin concentration to be relatively higher.

Based on the overall results, it can be concluded that all parameters and variables in the modified Hovorka equations took part in the simulation work. Hence, the interaction among the parameters and variables in glucose-insulin dynamic system was much better as compared to the original Hovorka equations. The simulation results showed that the blood glucose level was better controlled in its steady state when the modified Hovorka equations were applied in the simulation work. The modified Hovorka equations showed that the lower U_t at 16.7 mU/min and 20 mU/min could achieve a safe range of blood glucose level and lower plasma insulin concentration compared to the original Hovorka equations. For future work, it is recommended to conduct simulation works with meal disturbance using the modified Hovorka equations so as to observe its performance and applicability.

ACKNOWLDGEMENTS

The authors wish to acknowledge and extend their gratitude to the Malaysia Ministry of Higher Education (MOHE) and Institute of Research Management & Innovation (IRMI), Universiti Teknologi MARA (UiTM) for the financial support (600-RMI/DANA 5/3 BESTARI (060/2017). The authors would also like to thank the Faculty of Chemical Engineering and Accounting Research Institute (ARI), UiTM for all the support and assistance given.

REFERENCES

Bilous, R. W., Donnelly, R., & Williams, G. (2010). *Handbook of diabetes* (4th Ed.). Chichester, England: Wiley Blackwell.

Elleri, D., Dunger, D. B., & Hovorka, R. (2011). Closed-loop insulin delivery for treatment of type 1 diabetes. *BMC Medicine*, *9*(1), 1-9.

Forlenza, G. P., Buckingham, B., & Maahs, D. M. (2016). Progress in diabetes technology: developments in insulin pumps, continuous glucose monitors, and progress towards the artificial pancreas. *The Journal of Pediatrics*, *169*, 13-20.

Hovorka, R., Canonico, V., Chassin, L. J., Haueter, U., Massi-Benedetti, M., Federici, M. O., & Wilinska, M. E. (2004). Nonlinear model predictive control of glucose concentration in subjects with type 1 diabetes. *Physiological Measurement*, *25*(4), 905-920.

Hovorka, R., Wilinska, M. E., Chassin, L. J., & Dunger, D. B. (2006). Roadmap to the artificial pancreas. *Diabetes Research and Clinical Practice*, *74*, S178-S182.

Hovorka, R., Allen, J. M., Elleri, D., Chassin, L. J., Harris, J., Xing, D., & Dunger, D. B. (2010). Manual closed-loop insulin delivery in children and adolescents with type 1 diabetes: a phase 2 randomised crossover trial. *The Lancet*, *375*(9716), 743-751.

Messori, M., Cobelli, C., & Magni, L. (2015). Artificial pancreas: from in-silico to in-vivo. *IFAC-PapersOnLine*, *48*(8), 1300-1308.

Thabit, H., & Hovorka, R. (2012). Closed-loop insulin delivery in type 1 diabetes. *Endocrinology and Metabolism Clinics, 41*(1), 105-117.

Thabit, H., & Hovorka, R. (2016). Coming of age: the artificial pancreas for type 1 diabetes. *Diabetologia*, *59*(9), 1795-1805.

Yusof, N. F., Som, A. M., Ibrehem, A. S., & Ali, S. A. (2012, December 17-19). Parameter addition in interaction of glucose and insulin for type 1 diabetes. In *2012 IEEE EMBS Conference on Biomedical Engineering and Sciences (pp.* 273-278). Langkawi, Malaysia.

Yusof, N. F. M., Som, A. M., Ibrehem, A. S., & Ali, S. A. (2013). A review of mathematical model describing insulin delivery system for type 1 diabetes. *Journal of Applied Sciences, 14*, 1465-1468.

Yusof, N. F. B. M., Som, A. M., Ibrehem, A. S., & Ali, S. A., (2014). System identification in modified diabetic model for nanochip controller. *Advanced Materials Research, 938*, 299-304.

Yusof, N. F., Som, A. M., Ali, S. A., & Anuar, A. L. A. H. (2015). Simulation work for the control of blood glucose level in type 1 diabetes using hovorka equations. *Advanced Materials Research, 1113*, 739-744.

Pertanika J. Sci. & Technol. 27 (4): 1539 - 1559 (2019)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Composite Resin Veneer Systems: An In Vitro Study to Evaluate Surface Roughness Changes with Different Surface Treatments and Shear Bond Strength

Nik Zarina Nik Mahmood[1]*, Amar Mohammed Thiyab[1] and Mohamed Ibrahim Abu Hassan[2]

[1]Centre of Comprehensive Care Studies, Faculty of Dentistry, Universiti Teknologi MARA, 47000 UiTM, Sungai Buloh, Selangor, Malaysia
[2]Centre of Restorative Dentistry Studies, Faculty of Dentistry, Universiti Teknologi MARA, 47000 UiTM, Sungai Buloh, Selangor, Malaysia

ABSTRACT

This study aimed to evaluate surface roughness (SR) changes with surface treatments and shear bond strength (SBS) of two prefabricated and one laboratory-made composite veneer systems. The prefabricated groups, Edelweiss (EDL) and Componeer (CMP) while, SR Nexco (NEX) was a laboratory-made group. A total of hundred twenty samples, comprising 40 samples for each group were divided into four subgroups of surface treatment (n=10): (a) no treatment (control), (b) 9% hydrofluoric acid (HF), (c) abrasion with a high-speed diamond bur and (d) sandblast with aluminum trioxide (Al_2O_3) particles. A profilometer was used to evaluate the SR before and after surface treatments applications. Stereo electron microscope was utilized to assess changes occurred on the surface texture of the veneers. For SBS test, the prepared veneer was uploaded over an epoxy resin mould. Two cylindrical adhesive resins were bonded binary and perpendicular over the inner surface of the veneer and tested using Universal Testing Machine (SHIMADZUTM, Japan). Data was analyzed using One-way ANOVA, post-

ARTICLE INFO

Article history:
Received: 18 September 2018
Accepted: 17 May 2019
Published: 21 October 2019

E-mail addresses:
drnikzarina@uitm.edu.my (Nik Zarina Nik Mahmood)
amar6652@salam.uitm.edu.my (Amar Mohammed Thiyab)
mibrahim@salam.uitm.edu.my (Mohamed Ibrahim Abu Hassan)
**Corresponding author*

ISSN: 0128-7680
e-ISSN: 2231-8526

hoc student's t-test and Duncan test with $p<0.05$. One-way ANOVA revealed a significant increase in the SR of all veneer groups treated with a diamond bur and Al2O3 sandblast. NEX group showed higher SR (6.52 ± 0.85) followed by EDL (4.59 ± 0.75) and CMP (4.99 ± 0.67) groups. The significant higher SBS was demonstrated by NEX (22.88 ± 5.2 MPa). EDL exhibited higher bond strength (12.3 ± 3.7 MPa) than CMP (11.75 ± 6.5 MPa). A laboratory-made system produced higher SR enhancement with a diamond bur and superior bond strength.

Keywords: Composite veneer systems, surface roughness changes, surface treatments, shear bond strength

INTRODUCTION

Dental veneers are used to improve the color of discolored teeth and straightening slightly malposition teeth (Christensen, 2004). A dental veneer is defined as a layer of tooth-colored restorative material, usually porcelain or composite resin, attached to the surface by direct fusion, cementation or mechanical retention. Dental veneer systems can be classified according to its material and the mode of clinical usage. Christensen (2004) described the common types of veneer materials were porcelain and composite resin materials. In an earlier article, it was concluded that the mode of the clinical usage could be either direct; which most of the time used composite resin materials, or indirect when the veneer had to be manufactured in a laboratory before its clinical usage (Christensen, 2003). Composite resin veneer can also be classified as a laboratory-made system or prefabricated system (Toh et al., 1987).

Composite resin veneer is the technique of choice in complex rehabilitation cases as reported (Asensio-Acevedo et al., 2013). Indirect composite veneers were used to rehabilitate patients and have been widely used in an aesthetic restorative field due to its excellent properties such as wear resistance, aesthetics, marginal adaptation and enamel control over polymerization shrinkage, lower modulus elasticity and higher capacity to absorb functional stresses of composite restorations. Nandini (2010) reviewed several articles with regards to the improvement in properties of indirect resin composites and concluded that composite veneer could effectively be the alternative to the porcelain veneer in aesthetic dentistry.

The evolution of indirect composite veneer is primarily to improve its polymerization shrinkage which could lead to marginal microleakage, postoperative sensitivity, and recurrent caries (Loguercio et al., 2002). It is also used to supplement ceramic restorations in implant cases or patients with poor periodontal structures who require occlusal coverage. Leinfelder (2005) discovered composite veneer could absorb more occlusal stress compared to the ceramic material. The indirect composites are designed, modeled and cured extra-orally by dental technicians.

Prefabricated composite veneers are made of resin materials with specific polymerizing and finishing techniques. A review article has shown the advantages of composite resin materials are its aesthetic, high bond strength, and superior mechanical properties (Rosenstiel et al.,1998). However, Le Roux and Lachman (2007) concluded several shortcomings of these materials included wear, leakage and discoloration leading to impairment in the aesthetic value of the composite resin over time. Componeer (Coltène/Whaledent AG™, USA) and Elderweiss (Edelweiss, Ultradent Inc™, USA) are the two examples of prefabricated composite veneer systems available in the market. They are manufactured from a nanohybrid composite and are extremely thin veneer (0.3 mm) which allow conservation of tooth structure. Various studies have shown the micro-retentive inner surface ensures a lasting bonding, therefore conditioning of the veneer is not required studies (Gomes & Perdigão, 2014; Gurtu et al., 2016).

Surface roughness is a component of surface texture. It is quantified by the deviations in the direction of the normal vector of a real surface from its ideal form. If these deviations are large, the surface is rough; if they are small, the surface is smooth. A surface treatment is a process applied to the surface of a material to improve its retention form, for example by making it more resistant to corrosion or wear (Abu-Eittah, 2012), Barragan et al. (2014) studied different surface treatments effect on zirconia and composite resin and they found the most common pre-surface treatments used were with several mechanical and chemical agents such as aluminum oxide (Al_2O_3) sandblast, modified tribochemical technique or etching with hydrochloric acid (HF) and ferric chloride.

In a study to assess the different surface treatments effect between zirconia and dentin interface, it was concluded that the surface treatment, either for restorative material or the tooth surface, was considered one of the most common clinical procedures to increase the shear bond strength for the tooth-restoration complex (Abu-Eittah, 2012). With the application of surface treatments, the surface texture, or surface roughness, of the restoration will be increased. It can be described by the number of cavities, porosities, and grooves that have been made over the surface of restoration after treatment, thus, leading to the improvement of the quality of mechanical retention at the tooth-restoration complex. However, many factors deal with shear bond strength such as wettability of the bonding surface, silane application, micromechanical and chemical bonds have to be enhanced beside the surface roughness to achieve higher shear bond strength (Brosh et al., 1997; Schmidlin et al., 2010).

Several methods of surface treatments have been employed to roughening the inner surface of veneers mechanically and chemically. Air abrasion with 50 µm Al_2O_3 particles (Swift Jr et al., 1992), roughening with silicon carbide paper or diamond stones (Joulaei et al., 2012) and etching with HF are the examples of surface treatments available clinically. Other mechanical surfaces roughened like grinding with a diamond disc, abraded with a diamond bur or polishing with red a rubber wheel bur were rarely used (Mohammed et al., 2015).

Sandblasting is considered the most reliable mechanical surface roughened method that was used in these studies (Brosh et al., 1997; Grover & Nandlal, 2015). This method is used either with porcelain, zirconia or composite resin materials. Sandblasting is usually performed with aluminum powder with different sizes and time intervals (Grover & Nandlal, 2015; Su et al., 2015; Zhou et al., 2014).

Several studies utilized chemical agents to roughening a veneer surface (Fuentes et al., 2013; Poskus et al., 2015; Yavuz et al., 2013). Hydrofluoric acid, tribochemical silica coating system and phosphoric acid or salinized are few examples of chemical agents commonly used. Some of the studies used both mechanical and chemical roughening agents (Schmidlin et al., 2010; Zhou et al., 2014). Sandblast with 50µm Al_2O_3 particles is considered the most common mechanical surface treatment (Zhou et al. 2014). Schmidlin et al. (2010) stipulated that an abrasion of the surface with a high-speed diamond bur was considered the most commonly used in a clinical setting. HF with 9% or 12% concentration is commonly used for repairing the ceramic veneers with composite resin restoration. Other means of surface treatment are 98% sulfuric acid etching and Aragon plasma treatment (Ho et al. 2015).

Generally, maximum surface roughened increased the bond strength to the surface (Huang et al., 2013). But this concept has its own limitation, which is related to the restorative materials such as chemical compositions and mechanical properties of the material (Yenisey et al., 2016). This eventually leads to the changes from hydrophilic to hydrophobic properties during the setting of the composite resin cement.

Most of the studies evaluated shear bond strength of porcelain on alloy and tooth structures such as enamel and dentin (Çiftçi et al., 2007). They reported porcelain fused to metal (PFM) showed considreably higher shear bond strength than with an adhesive resin. Another study has experimented their self adhesive resin cement to conventional composite resin cement such as Variolink, Panavia F2.0, RelyX Unicem and Maxcem Elite. The authors found that the shear bond strength of self-adhesive resin cement was inferior compared to conventinal cements (Lührs et al., 2010). Until now, no studies have been investigated the shear bond strength between different types of composite veneer systems.

To date, not many studies have been conducted to evaluate the effect of different surface treatments on the surface roughness and bond strength of prefabricated and laboratory-made composite veneers. Therefore, this study is aimed to evaluate the surface roughness (SR) changes with different types of surface treatment and bond strength between two prefabricated and one laboratory-made veneer systems.

The null hypotheses of this study are (1) type of material and (2) type of surface treatment would demonstrate no difference on the surface roughness and bond strength of composite resin veneer materials.

MATERIALS AND METHODS

This is a laboratory-based experimental study evaluating and comparing the surface roughness (SR) changes of two prefabricated and one laboratory-made veneer systems with different surface treatments. The prefabricated systems were Group EDL (Edelweiss, Ultradent Inc™, USA) and Group CMP (Coltène/Whaledent AG™, USA) while the laboratory-made system was Group NEX (SR Nexco, Ivoclar Vivadent, Liechtenstein).

A total of hundred twenty samples (n=120) were tested in this study. Forty veneers were selected from each group: EDL (n=40), CMP (n=40) and NEX (n=40). Each group was further were assigned to 4 subgroups according to surface treatment methods (n=10):

Group 1: no treatment
Group 2: an abrasion with a high-speed diamond bur
Group 3: 9% hydrofluoric acid (HF)
Group 4: sandblast with aluminum trioxide (Al2O3) particles.

The sample size was estimated based on previous literature (Schmidlin et al., 2010; Swift Jr et al., 1992). All samples were cleaned by an ultrasonic water bath (Renfert SYMBRO, Germany) with plaster solvent chemical solution (Gypsolve, England) to remove any accretion or industrial smear layer. Following that, all samples were dried with non-oily dry air and stored in containers. These samples were treated with four different surface treatments simultaneously.

Surface Treatments

No treatment (Group 1/ control group). Ten veneers from each sample group were randomly selected and kept in the container contains saline water without any treatment and served as a control group.

Abrasion with a high-speed diamond bur (Group 2). Ten veneers from each sample group were abraded with a high-speed oval diamond bur by using turbine hand-piece at a speed of 160,000 rpm. The abrasion area was created in three different points on the inner surface of the veneer. These points were positioned at cervical, medial and incisal with 2 mm distance apart (Figure 1). The procedure was done manually using x 3.0 Galilean dental surgical loupes (Univet Optical Technologies North America Inc., Markham, Canada) to simulate the clinical technique. The position of a diamond bur was parallel to the inner surface with a minimal force for one second at each point.

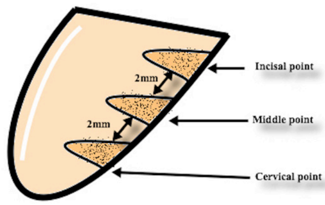


Figure 1. Three different abrasion points on the inner surface of the veneer.

9% Hydrofluoric Acid (HF) Treatment (Group 3). Ten veneers from each sample group were randomly selected and arranged in a line to be treated with 9% HF. By using a micro brush, hydrofluoric acid was applied on the inner surface of each sample for 60 seconds, washed with water spray and dried with non-oily air spray as recommended by manufacturers. Then, all samples were transferred to the labeled containers according to its groups using a plastic tweezer and clinical gloves to avoid contamination before surface roughness measurement.

Sandblast with Aluminum Trioxide (Al2O3) particles (Group 4). Ten veneers from each sample group were randomly selected and sandblasted with aluminum trioxide (Al_2O_3) particles by using sandblast machine (Duostar BEGOTM, Germany) under the pressure of 2 bars, a 10mm distance between the sample and the sandblast tube with 5 seconds exposure time. The sandblast tube was positioned perpendicular over the inner surface of each sample. All samples were cleaned under running water and dried with non-oil air spray to remove any remaining foreign bodies. They were stored in the containers and before surface roughness measurement with a contact profilometer (AMBIOS XP-1, USA).

Surface Roughness (Ra) Evaluation

All samples were transferred to the stage of AMBIOS XP-1 to quantify the Ra value at the three points positions. The mean reading of each veneer was considered as the main Ra value. The profilometer scan was set-up at an equal speed to 0.05mm/sec, 1.5mm length, 10 microns range, 1.0 mg stylus range, filter level equal to 5 and irregular surface as profile type. This set-up was used for all four test groups. The measurement of each sample was done using gloves and a plastic tweezer to prevent any surface contamination. The Ra value was recorded by a device software and the data was transferred to Microsoft Excel sheet 2013.

Scanning Electron Microscope (SEM) Evaluation

After completed of surface roughness testing, each sample from the sample group was selected and observed under Scanning Electron Microscope (SEM) with the magnification powers of 100x and 1000x. The purpose of this procedure was to assess any changes in the surface texture before and after application of the surface treatments. The selected samples were coated with Sputter coater (EdwardsTM, USA) before covering it with a thin layer of conducting material (silver alloy). A conductive coating is required to prevent charging of specimens with an electron beam in conventional SEM mode of high vacuum and voltage.

Shear Bond Strength Testing

The study was further elucidated to investigate which veneer groups exhibited the highest shear bond strength. Samples were carefully selected and were utilized from each veneer system (n=10); CMP, EDL, and NEX, that had undergone surface treatment and showed the superior Ra values. The veneer was trimmed with diamond saw under cooling water to produce a flat bonding surface and then cut into a rectangular shape measured 8mm length with 5mm width. The veneer was placed horizontally on a translucence rectangular box filled with epoxy resin (Maricon, Romania) with the inner surface of each veneer was positioned away from the epoxy resin.

Prior to the placement of an adhesive materials, the veneer surface was etched for 20 seconds to remove any debris, washed under running water and dried using air spray. Two mould frames, each measured 2.5mm in diameter and 1.5mm thickness were created by a specific leather puncture and an acid-resistant adhesive tape were located on the inner surface of the veneer. The frames were loaded with an adhesive agent, light cured according to the manufacturer's guide to produce two adhesive cylinders. A magnifying dental loupe was used to check any defect on the adhesive cylinders and were discarded if any defect found. The calculation of sample size for SBS testing was based on the previous studies (n=10*2=20) (Abo-Hamar, 2013; Khamverdi et al., 2013; Pahlavan et al., 2013; Perdigão et al., 2013; Puvoravan et al., 2013). Shear bond strength was determined by Universal Testing Machine (SHIMADZU™, Japan). The shear bond set-up pin of the machine was perpendicularly positioned over the cylinder and parallel to the veneer inner surface. A shear load was applied at a crosshead speed of 1 mm/min until its failure. The micromechanical bond strength value is indicated in Figure 2.

Statistical Analysis

The data was recorded by the Statistical Package for the Social Sciences (SPSS) version 17.0 (SPSS Inc. Chicago, Illinois, USA). Descriptive statistics including means and standard deviations were calculated for each group, compared and analysed using One-way analysis of variance (ANOVA) to compare the significance difference of surface roughness and shear bond strength between groups. Post-hoc, student t-test was used to determine which veneer group showed the highest Ra value. Two-way ANOVA was employed to analyze Ra value which comparing between two dependent variables (three veneer groups and four surface treatments). Student t-test was also used to evaluate the most superior shear bond strength among the veneer systems. A statistically significance difference was determined at a 95% confidence level with a p-value of ≤ 0.05.

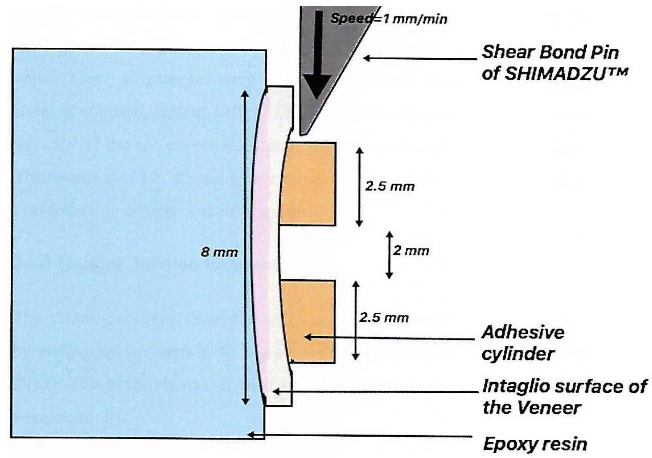


Figure 2. The diagram showed the position of adhesive agent cylinders in relation to shear bond pin set-up.

RESULTS

Three methods of comparison were conducted in this study. The first comparison was done between the sample groups without any application of surface treatments with the aimed to evaluate any significance of Ra value. Figure 3 as indicated by Group 1 shows the Ra values of all samples prior to surface treatments. Descriptive analysis demonstrated the highest Ra value came from Group CMP (1.58 ± 0.66) followed by Group EDL (0.57 ± 0.25) and Group NEX (0.38 ± 0.18).

The second comparison was when the sample groups were treated with the tested surface treatments. The objective was to evaluate any significant difference of the Ra value upon receiving three different methods of surface treatments. Bar chart in Figure 3 shows the highest Ra value was from Group 2 with Group NEX (6.52 ± 1.20) followed by Group CMP (4.99 ± 0.94) and Group EDL (4.59 ± 1.05). Meanwhile, the lowest Ra value was from hydrofluoric acid surface treatment with Group NEX (0.38 ± 0.12), Group EDL (0.97 ± 0.67) and Group CMP (1.36 ± 0.32). Among the sample groups, Group NEX exhibited the highest Ra value followed by Group CMP and Group EDL.

Based on the outcome of the surface treatment methods, air abrasion with an oval-shaped high-speed diamond bur exhibited the highest Ra value. Therefore, the third objective is to use this method for further investigation the SBS among the veneer groups. NEX group showed the highest SBS (22.88 ± 5.2 MPa) as compared to prefabricated veneer groups; EDL (12.31 ± 3.7 MPa) and CMP (11.75 ± 6.5 MPa) with p-value <0.05. There was no significant difference between the EDL group and the CMP group, p<0.05.

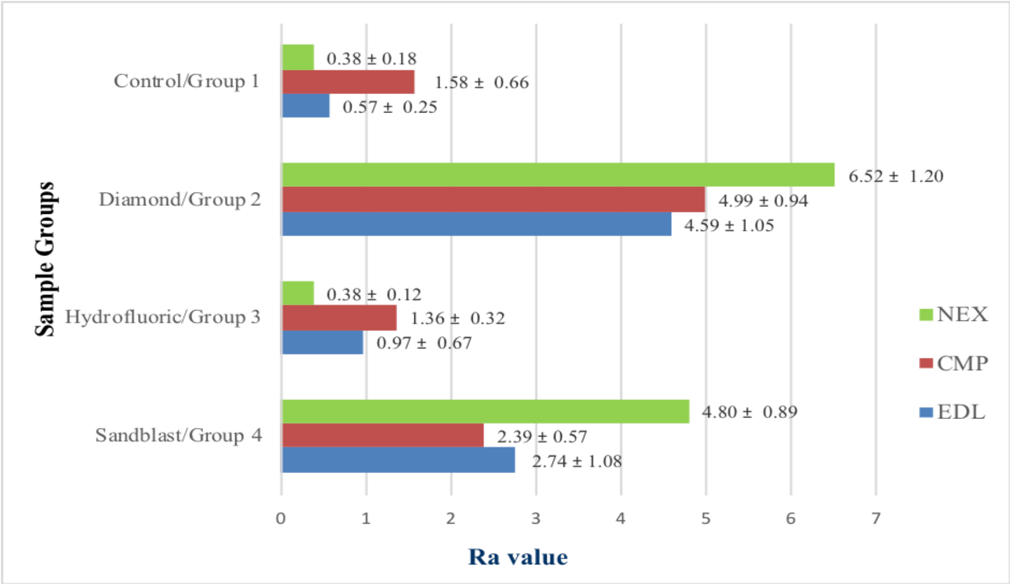


Figure 3. The mean values and standard deviation of surface roughness (Ra) for each surface treatment of the three sample groups.

Statistical Analysis with One-way ANOVA and Student t-test

Non-treatment group (control group). One-way ANOVA exhibited a significant difference for all Ra values between samples in Group 1 (Table 1). Table 2 shows the significant differences between Group CMP with Group EDL and Group CMP with Group NEX ($p<0.05$). No significant difference was recorded between Group EDL with Group NEX.

Surface Treatment groups (test groups). Irrespective of the sample groups, one-way ANOVA revealed a significant difference of Ra values when the sample surface was treated by the three surface treatments (Table 3). Student's t-test was carried out for each surface treatment method with the sample groups. Table 2 displays a significant difference of SR value by Group 2 between Group NEX with Group EDL ($p=0.01$) and Group CMP ($p=0.05$). Nevertheless, no significant difference was found between Group EDL with Group CMP ($p=0.389$) regardless of the surface treatment methods. Table 2 shows similar significance results for Group 3 and Group 4 respectively ($p<0.05$).

Two-way ANOVA exhibited no significant difference of Ra value by NEX and CMP veneer systems prior to the treatment or after treatment with Group 3. The same result was found with NEX and EDL systems, no significant difference in Ra value was noted prior to the treatment and after treatment with Group 3. Group 2 and Group 4 demonstrated almost similar surface roughness for NEX and EDL veneer systems. Generally, it was found that both prefabricated and a laboratory-made group showed significantly higher RA value with Group 2 and Group 4 but not with Group 3 (Table 4) ($p<0.05$).

Table 1

One-way ANOVA for the surface roughness values on non- treated test groups

Surface treatment methods	St. Error	Significant F>C	P-value
Group 1	0.1	Yes	0.000

Note. Average surface roughness (Ra) varies across values of non-treatment and the three treatment veneer groups (α =0.05).

Table 2

Student's t- test for surface roughness values between surface treatment methods with the three veneer groups

Surface Treatment Methods	Sample Groups		St. Error	Significant t>C	P-value
Group 1	CMP	EDL	0.2	Yes	0.000
		NEX	0.2	Yes	0.000
	EDL	CMP	0.2	Yes	0.000
		NEX	0.1	No	0.068
	NEX	CMP	0.2	Yes	0.000
		EDL	0.1	No	0.068
Group 2	CMP	EDL	0.2	No	0.389
		NEX	0.3	Yes	0.005
	EDL	CMP	0.2	No	0.389
		NEX	0.3	Yes	0.001
	NEX	CMP	0.3	Yes	0.005
		EDL	0.3	Yes	0.001
Group 3	CMP	EDL	0.1	No	0.113
		NEX	0.1	Yes	0.000
	EDL	CMP	0.1	No	0.113
		NEX	0.1	Yes	0.000
	NEX	CMP	0.1	Yes	0.000
		EDL	0.1	Yes	0.000
Group 4	CMP	EDL	0.2	No	0.375
		NEX	0.3	Yes	0.000
	EDL	CMP	0.2	No	0.375
		NEX	0.3	Yes	0.000
	NEX	CMP	0.3	Yes	0.000
		EDL	0.3	Yes	0.000

Table 3
One-way ANOVA for the surface roughness values by each surface treatment method on the test veneer groups

Surface Treatments Methods	St. Error	Significant F>C	P-value
Group 2	0.2	Yes	0.000
Group 3	0.1	Yes	0.000
Group 4	0.3	Yes	0.000

Note. Average surface roughness (Ra) varies across values of non-treatment and the three treatment veneer groups (α =0.00)

Table 4
Two-way ANOVA analysis for surface roughness values and surface treatments between a laboratory-made and two prefabricated veneer groups

G1	G2	Surface treatment		St. Error	Significant t>C	P-value
NEX	CMP	Group 1	Group 2	0.4	Yes	0.000
		Group 1	Group 3	0.1	No	0.374
		Group 1	Group 4	0.3	Yes	0.000
		Group 2	Group 3	0.4	Yes	0.000
		Group 2	Group 4	0.3	No	0.143
		Group 4	Group 3	0.3	Yes	0.000
	EDL	Group 1	Group 2	0.4	Yes	0.000
		Group 1	Group 3	0.1	No	0.103
		Group 1	Group 4	0.3	Yes	0.000
		Group 2	Group 3	0.4	Yes	0.000
		Group 2	Group 4	0.3	No	0.844
		Group 4	Group 3	0.3	Yes	0.000

Scanning Electron Microscope (SEM)

The SEM was further used to ascertain the changes in the surface texture of the inner surface for each sample group after various surface treatments. Under 1000x magnification power, the surface texture of a randomly selected sample from each sample group was evaluated and compared between treated and non-treated groups as depicted in Figure 4, 5 and 6.

It was clearly observed that oblique stripes, pores, and cavities were established on the inner surface of the sample when treated with Group 2 and Group 4 methods as indicated by a black arrow in Figure 4b, 4d, 5b, 5d, 6b and 6d. In contrast, the Group 3 surface treatment demonstrated a relatively smooth surface and minimum irregularities on the inner surface of all sample groups (a black arrow in Figure 4c, 5c, 6c).

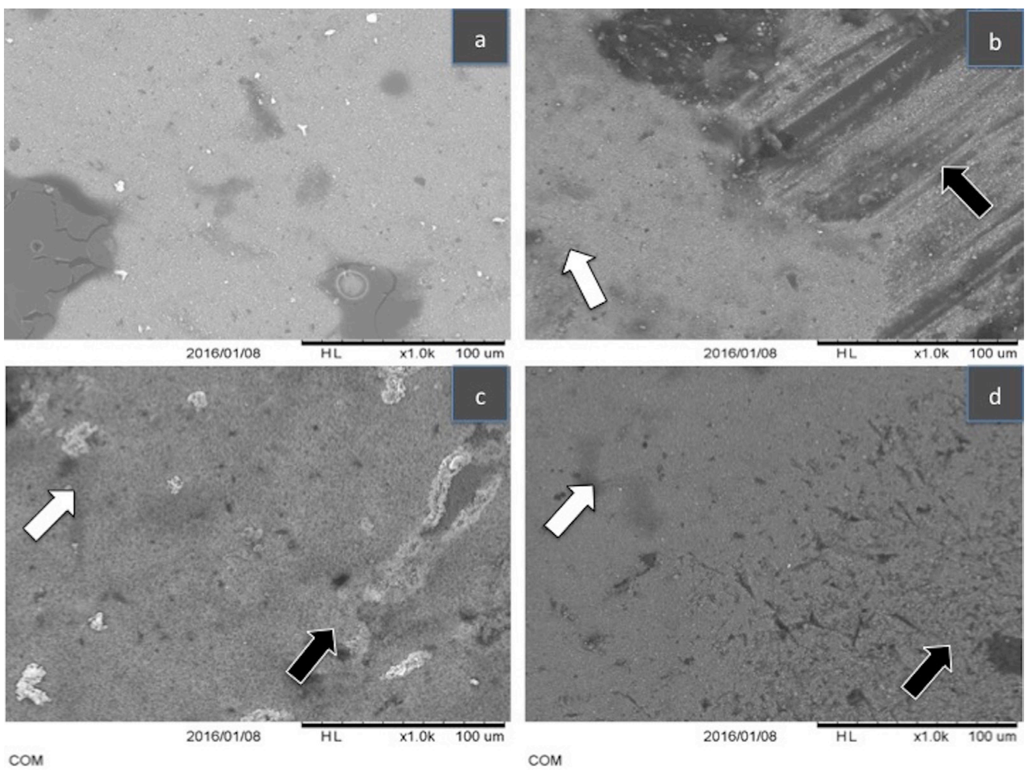


Figure 4. SEM images of the inner surface of CMP veneer before and after various surface treatments.

a) Group 1 (control), no treatment
b) Group 2 (abrasion with a high-speed diamond bur)
c) Group 3 (hydrofluoric acid (HF) treatment) presented with a smooth surface and minimum irregularities
d) Group 4 (sandblast with Aluminum Trioxide (Al2O3) particles)

The white arrow indicates the image before surface treatment and the black arrow indicates after various surface treatment

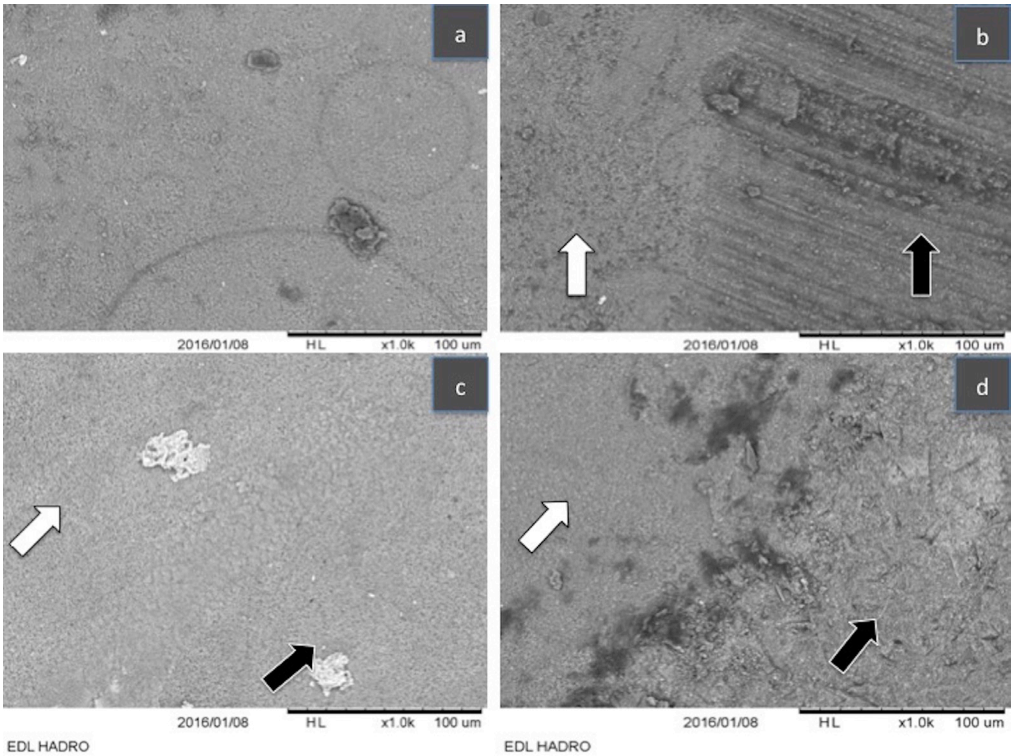


Figure 5. SEM images of the inner surface of EDL veneer before and after various surface treatments.
a) Group 1 (control), no treatment
b) Group 2 (abrasion with a high-speed diamond bur)
c) Group 3 (hydrofluoric acid (HF) treatment) presented with a smooth surface and minimum irregularities
d) Group 4 (sandblast with Aluminum Trioxide (Al2O3) particles)

The white arrow indicates the image before surface treatment and the black arrow indicates after various surface treatment

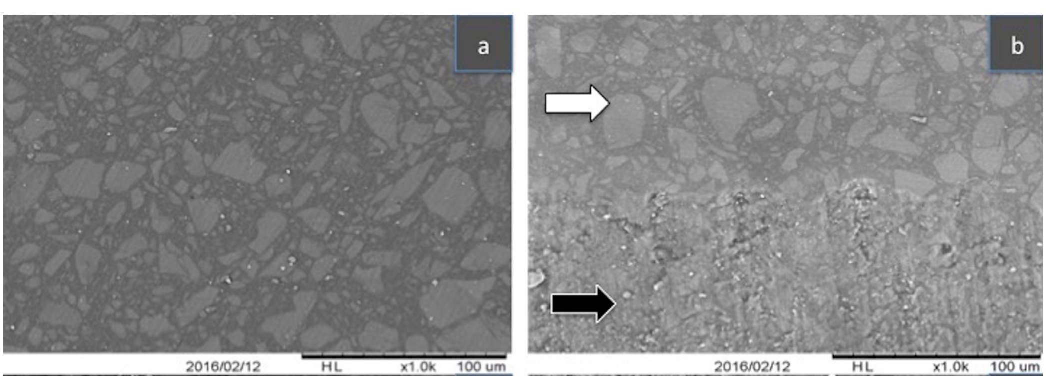


Figure 6. The SEM images of the inner surface NEX veneer before and after the surface treatments.
a) Group 1 (control), no treatment
b) Group 2 (abrasion with a high-speed diamond bur)

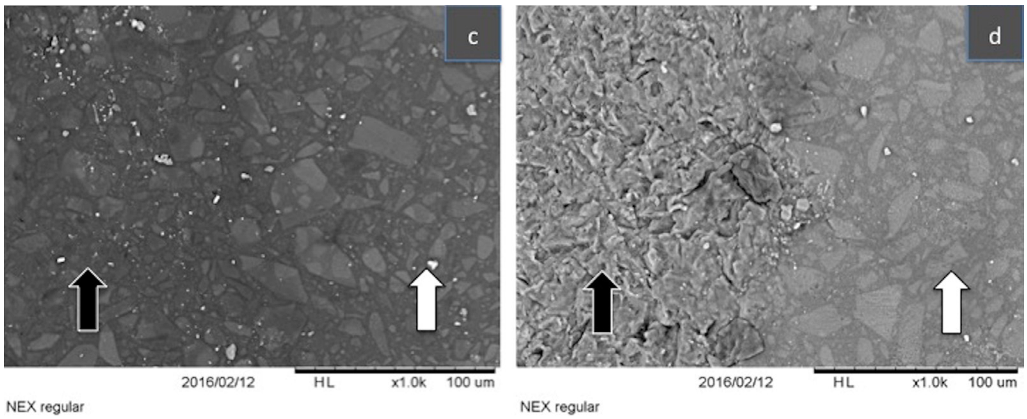


Figure 6 *(Continued)*
c) Group 3 (hydrofluoric acid (HF) treatment) presented with a smooth surface and minimum irregularities
d) Group 4 (sandblast with Aluminum Trioxide (Al2O3) particles)

The white arrow indicates the image before surface treatment and the black arrow indicates after various surface treatment

Table 5
Student t- test for shear bond strength values within veneer groups

		Test Statistic t	Critical Value @5 % C	St. Error	Significant t>C	P-value
CMP	EDL	0.302	2.042	0.9	No	0.765
	NEX	5.515	2.042	1.4	Yes	0.000
EDL	CMP	0.302	2.042	0.9	No	0.765
	NEX	6.803	2.042	1.2	Yes	0.000
	CMP	5.515	2.042	1.4	Yes	0.000

Statistical Analysis with One-way ANOVA and Student t-test

One-way ANOVA revealed a significance different of SBS was found between 2 types of veneer systems. Table 5 shows the significant differences between Group CMP and Group EDL with Group NEX ($p=0.00$). No significant difference was recorded between Group CMP and Group EDL ($p>0.05$)

DISCUSSION

This is an *in-vitro* study focusing on the surface roughness changes of two types of composite resin veneer systems with three different surface treatments. The surface roughness is important to ensure a good mechanical pre-treatment bonding between restorative materials and adhesive systems.

Surface treatments were applied to the inner surface of these samples with the aimed to investigate the best method to enhance its mechanical or chemical retention, which is indicated by the increased Ra value. Surface treatments included an abrasion with a high-speed diamond bur, acid etching with 9% HF and sandblasting with 50 um aluminum trioxide (Al_2O_3) particles were tested. A descriptive analysis was carried out on the control group prior to the surface treatments applications primarily to assess the pre-treatment Ra value for each veneer group. Based on Figure 3 (Group 1), it could be suggested that CMP veneers demonstrated the roughest surface prior to its treatments.

One-way ANOVA showed significant differences of Ra values by three sample groups when treated with various types of surface treatments. Comparing the materials within the same surface treatment methods revealed that Group NEX generally demonstrated the highest Ra value especially when treated with Group 2. It could be suggested that a laboratory-made veneer system has superior Ra value than prefabricated veneer systems. With regards to the prefabricated veneers, students t-test analysis revealed no significant difference of Ra value was found between Group CMP and Group EDL. This indicates that both groups have almost the same quality of surface roughness irrespective of the methods of surface treatments.

This study also demonstrated the application of surface treatments significantly increased the quality of surface roughness of composite resin veneers when compared to a control group (Figure 3, Table 4). Güngör et al. (2016) and Neis et al. (2015) stipulated an increased surface roughness improves mechanical interlocking on the bonding surface and its bond strength on dental ceramics. Surface treatments have been shown to improve the bond strength of resin composite to CAD/CAM resin-ceramic hybrid materials for repair (Elsaka, 2015; Stawarczyk et al., 2015; Wiegand et al., 2015). Grinding with a diamond bur generated the highest surface roughness among the surface treatment methods on resin composite bonded to resin- ceramic hybrid materials (Güngör et al., 2016). The result of the present study is in accordance with those studies.

Two-way ANOVA revealed a significant increase in Ra value as showed by Group 2 and Group 4 surface treatments ($p<0.05$). It can be concluded that both methods were effective to increase the mechanical retention form for all sample groups (Table 4). However, it was a clear evidence (Roeters, 2000) that a major disadvantage of sandblasting was the aerosol of fine abrasive particles that will contaminate a wide area of the operatory, which might be harmful to patients and operators. Therefore, it is recommended to abrade the inner surface of veneer with a diamond bur as a surface pre-treatment prior to cementing process to obtain a stronger mechanical bond between the two structures. Moreover, this method can be easily applied and produce a faster result clinically. Nevertheless, an air abrasion with silica and (Al_2O_3) sandblast for 15s at a distance of 10 mm has been shown to increase surface roughness with PEEK composite resins during composite repairs (Poskus et al., 2015; Zhou et al., 2014).

In this study, 9% hydrofluoric acid produced the lowest Ra value for the tested samples. No significant difference was exhibited between Group 3 and Group 1. It can be extrapolated that 9% HF surface treatment is not chemically effective in improving the surface roughness of the composite veneers. This finding is consistent with previous studies (Swift Jr et al., 1992; Zhou et al., 2014).

The sample surfaces treated with hydrofluoric acid yielded the lowest RA value. Therefore, it could be suggested that the hydrofluoric acid is not recommended to be used as a routine pre-treatment method. Furthermore, hydrofluoric acid has a corrosive effect and a contact poison. A meticulous application technique is needed to prevent detrimental side effects, such as acid burns and necrosis of the underlying soft tissues as previously reported (Asvesti et al., 1997).

It was clearly evidenced under SEM with 1000x magnification, that oblique stripes, pores, and cavities were established on the inner surface of the sample when treated with Group 2 and Group 4 methods. This is shown by a black arrow in Figure 4b, Figure 4d, Figure 5b, Figure 5d, Figure 6b, Figure 6d compared to a white arrow (before the treatments) in the same figures. In contrast, Group 3 surface treatment demonstrated a relatively smooth surface and minimum irregularities on the inner surface of all sample groups as demonstrated by a black arrow in Figure 4c, Figure 5c, Figure 6c than a white arrow (before the application of Group 3). Their surface textures showed the same appearance towards different surface treatments. The differences in the size, volume, and shape of the fillers of the three composite veneer systems possibly did not significantly affect their surface roughness. This finding is concurrent with previous studies (Ho et al., 2015; Schmidlin et al., 2010).

A study has shown that the surfaces treated with hydrofluoric acid showed undercuts when the filler were dissolved and removed from the matrix (Zhou et al., 2014). This finding was quite similar to the present study even with a different concentration of the hydrofluoric acid. The weak bond strength of the HF on the pre-treated restorations could be attributed by the low surface roughness after treatment as proven in the present study.

The veneer samples which were treated by Group 2 were chosen to undergo shear bond strength test (SBS). The objective was to investigate which type of veneer systems that produced superior micromechanical bonding with its adhesive agent. According to ISO 10477 requirements, the minimum acceptable shear bond strength value is at least 5 MPa (Sarafianou et al., 2008). Matsumura et al. (2001) suggested the minimum of 10 MPa of SBS for resin to metal bonded had to be achieved to consider as clinical satisfactory. But, no literatures had been mentioned on the SBS for resin to resin.

The adhesive systems used for both prefabricated veneers as recommended by the manufacturer, Multilink Automix (Ivoclar Vivadent, USA) were selected to be used as an adhesive system on the laboratory-made veneer system. Völkel (2004) postulated the value of SBS for this adhesive agent was between 20-25 MPa as compared to the other agent such the self-curing RelyX Unicem (13.08 ± 3.61 MPa) (Moghaddas et al., 2017). With this evidence, we chose Multilink Automix adhesive agent due higher bond strength on both enamel and dentin structures. However, until now, no study has been conducted to evaluate SBS from resin to resin. In this study, the SBS for NEX group was 22.8 ± 5.0 MPa followed by EDL group; 12.3 ± 3.7 MPa and CMP group; 11.75 ± 6.5 MPa. One way ANOVA showed NEX group was significantly higher than the other two prefabricated groups. The significant difference could be attributed to many factors such as the type of adhesive system for each group and the size of fillers in veneer resin material itself. Perdigão et al. (2013) evaluated the micro shear bond strength of the prefabricated veneer system (Componeer™) and compared it with IPS and Cerinate (Ceramic based veneer). He found that the μSBS of Componeer is 15.2 ± 2.5 MPa which relatively is the same in the present study.

Several studies reported volume, weight, types and the size of fillers affects shear bond strength on composite resin (Gallo et al., 2001; Kim et al., 2002; Miyazaki et al., 1995). Further studies need to be carried out on shear bond strength of other methods of surface treatment such as sandblast with Al2O3 particles or 9% HF treatment. The composition of a laboratory-made veneer should interestingly be investigated to indicate its mechanical properties.

Few limitations have been discovered in this study such as limited resources of technical information related to the prefabricated veneer systems, thus, it led to the restriction in comparing the outcome of the present study. The future study also needs to take into consideration of the cost involved if it cracks during the preparation of the samples.

CONCLUSION

Based on the result of this study and within the limitations described, a laboratory-made veneer system had exhibited superior surface roughness than prefabricated veneer systems when treated with different types of surface treatments. Group NEX demonstrated the highest surface roughness value followed by Group CMP and Group EDL. Among the surface treatments tested, an abrasion with a high-speed diamond bur showed the best method to roughening the inner surface of composite veneers followed by sandblasting with aluminum trioxide (Al_2O_3) particles and 9% HF. This study also showed 9% HF agent did not produce any significant effect on the surface roughness and almost similar texture with the non-treated surfaces. The laboratory-made veneer system achieved the highest mean shear bond strength. EDL and CMP veneer systems appeared to be almost similar bond strength and it conveyed no statistical advantage between each other.

ACKNOWLEDGEMENT

We would like to express our gratitude to research laboratory officers at Faculty of Dentistry of University Teknologi MARA, Malaysia for their continuous support of the experimental devices for this research.

REFERENCES

Abo-Hamar, S. E. (2013). Effect of endodontic irrigation and dressing procedures on the shear bond strength of composite to coronal dentin. *Journal of Advanced Research*, *4*(1), 61-67.

Abu-Eittah, M. R. (2012). Assessment of different surface treatments effect on surface roughness of zirconia and its shear bond strength to human dentin. *Life Science Journal, 9*(4), 1792-1803.

Asensio-Acevedo, R., María Suarez-Feito, J., Suárez-Tuero, C., Jané, L., & Roig, M. (2013). The use of indirect composite veneers to rehabilitate patients with dental erosion: A case report. *European Journal of Esthetic Dentistry, 8*(3), 414-431.

Asvesti, C., Guadagni, F., Anastasiadis, G., Zakopoulou, N., Danopoulou, I., & Zographakis, L. (1997). Hydrofluoric acid burns. *Cutis, 59*(6), 306-308.

Barragan, G., Chasqueira, F., Arantes-Oliveira, S., & Portugal, J. (2014). Ceramic repair: Influence of chemical and mechanical surface conditioning on adhesion to zirconia. *Oral Health and Dental Management, 13*(2), 155-158.

Brosh, T., Pilo, R., Bichacho, N., & Blutstein, R. (1997). Effect of combinations of surface treatments and bonding agents on the bond strength of repaired composites. *Journal of Prosthetic Dentistry, 77*(2), 122-126.

Christensen, G. J. (2003). Direct restorative materials: What goes where? *The Journal of the American Dental Association, 134*(10), 1395-1397.

Christensen, G. J. (2004). What is a veneer?: Resolving the confusion. *The Journal of the American Dental Association, 135*(11), 1574-1576.

Çiftçi, Y., Canay, Ş., & Hersek, N. (2007). Shear bond strength evaluation of different veneering systems on ni-cr alloys. *Journal of Prosthodontics, 16*(1), 31-36.

Elsaka, S. E. (2015). Repair bond strength of resin composite to a novel cad/cam hybrid ceramic using different repair systems. *Dental Materials Journal, 34*(2), 161-167.

Fuentes, M. V., Ceballos, L., & González-López, S. (2013). Bond strength of self-adhesive resin cements to different treated indirect composites. *Clinical Oral Investigations, 17*(3), 717-724.

Gallo, J., Comeaux, R., Haines, B., Xu, X., & Burgess, J. (2001). Shear bond strength of four filled dentin bonding systems. *Operative Dentistry, 26*(1), 44-47.

Gomes, G., & Perdigão, J. (2014). Prefabricated composite resin veneers–a clinical review. *Journal of Esthetic and Restorative Dentistry, 26*(5), 302-313.

Grover, N., & Nandlal, B. (2015). An in vitro evaluation of the effect of sandblasting and laser surface treatment on the shear bond strength of a composite resin to the facial surface of primary anterior stainless steel crowns. *Journal of Clinical and Experimental Dentistry, 7*(1), e119-e125.

Güngör, M. B., Nemli, S. K., Bal, B. T., Ünver, S., & Doğan, A. (2016). Effect of surface treatments on shear bond strength of resin composite bonded to cad/cam resin-ceramic hybrid materials. *The Journal of Advanced Prosthodontics, 8*(4), 259-266.

Gurtu, A., Bansal, R., Chowdhary, P., Mehrotra, N., Mohan, S., Mehrotra, A., & Kishore, A. (2016). Componeer: An emerging esthetic solution. *Journal of Dental Sciences and Oral Rehabilitation, 7*(3), 146-148.

Ho, B. J., Tsoi, J. K. H., Liu, D., Lung, C. Y. K., Wong, H. M., & Matinlinna, J. P. (2015). Effects of sandblasting distance and angles on resin cement bonding to zirconia and titanium. *International Journal of Adhesion and Adhesives, 62*, 25-31.

Huang, B. R., Wang, X. Y., & Gao, X. J. (2013). Effects of different surface treatments on ceramic repairs with composite. *The Chinese Journal of Dental Research, 16*(2), 111-117.

Joulaei, M., Bahari, M., Ahmadi, A., & Oskoee, S. S. (2012). Effect of different surface treatments on repair micro-shear bond strength of silica-and zirconia-filled composite resins. *Journal of Dental Research, Dental Clinics, Dental Prospects, 6*(4), 131-137.

Khamverdi, Z., Rezaei-Soufi, L., Kasraei, S., Ronasi, N., & Rostami, S. (2013). Effect of epigallocatechin gallate on shear bond strength of composite resin to bleached enamel: An in vitro study. *Restorative Dentistry and Endodontics, 38*(4), 241-247.

Kim, K. H., Ong, J. L., & Okuno, O. (2002). The effect of filler loading and morphology on the mechanical properties of contemporary composites. *The Journal of Prosthetic Dentistry, 87*(6), 642-649.

Leinfelder, K. (2005). Indirect posterior composite resins. *Compendium of Continuing Education in Dentistry, 26*(7), 495-503

Le Roux, A. R., & Lachman, N. (2007). Dental composite materials: Highlighting the problem of wear for posterior restorations. *South African Dental Journal, 62*(8), 342-344.

Loguercio, A. D., Reis, A., Mazzocco, K. C., Dias, A. L., Busato, A. L. S., da Motta Singer, J., & Rosa, P. (2002). Microleakage in class ii composite resin restorations: Total bonding and open sandwich technique. *Journal of Adhesive Dentistry, 4*(2), 137-144.

Lührs, A. K., Guhr, S., Günay, H., & Geurtsen, W. (2010). Shear bond strength of self-adhesive resins compared to resin cements with etch and rinse adhesives to enamel and dentin in vitro. *Clinical Oral Investigations, 14*(2), 193-199.

Matsumura, H., Yanagida, H., Tanoue, N., Atsuta, M., & Shimoe, S. (2001). Shear bond strength of resin composite veneering material to gold alloy with varying metal surface preparations. *The Journal of Prosthetic Dentistry, 86*(3), 315-319.

Miyazaki, M., Ando, S., Hinoura, K., Onose, H., & Moore, B. K. (1995). Influence of filler addition to bonding agents on shear bond strength to bovine dentin. *Dental Materials, 11*(4), 234-238.

Moghaddas, M. J., Hossainipour, Z., Majidinia, S., & Ojrati, N. (2017). Comparison of the shear bond strength of self-adhesive resin cements to enamel and dentin with different protocol of application. *Electronic Physician, 9*(8), 4985-4991.

Mohammed, B., Afram, B., & Nazar, Z. (2015). An evaluation of the effect of different surface treatment on hardness and smoothness of pressable ceramic (in vitro study). *IOSR Journal of Dental and Medical Sciences, 14*(2), 84-89.

Nandini, S. (2010). Indirect resin composites. *Journal of Conservative Dentistry: JCD, 13*(4), 184-194.

Neis, C. A., Albuquerque, N. L. G., Albuquerque, I. D. S., Gomes, E. A., Souza-Filho, C. B. D., Feitosa, V. P., … & Bacchi, A. (2015). Surface treatments for repair of feldspathic, leucite-and lithium disilicate-reinforced glass ceramics using composite resin. *Brazilian Dental Journal, 26*(2), 152-155.

Pahlavan, A., Mehmanchi, M., Omrani, L. R., & Chiniforush, N. (2013). Effect of air abrasion and erbium-doped yttrium aluminum garnet (Er: YAG) laser preparation on shear bond strength of composite to dentin. *Journal of Lasers in Medical Sciences, 4*(3), 127-130.

Perdigão, J., Sezinando, A., Muñoz, M., Luque-Martinez, I., & Loguercio, A. (2013). Prefabricated veneers-bond strengths and ultramorphological analyses. *The Journal of Adhesive Dentistry, 16*(2), 137-46.

Poskus, L. T., Meirelles, R. S., Schuina, V. B., Ferreira, L. M., da Silva, E. M., & Guimarães, J. G. A. (2015). Effects of different surface treatments on bond strength of an indirect composite to bovine dentin. *Indian Journal of Dental Research, 26*(3), 289-294.

Puvoravan, C., Tanchanapradit, R., & Rachawang, N. (2013). Effect of different cleansing techniques on the shear bond strength between resin composite and dentin contaminated with eugenol-containing sealer. *Chulalongkorn University Dental Journal, 34*(3), 203-212.

Roeters, J. J. (2000). A simple method to protect patient and environment when using sandblasting for intraoral repair. *Journal of Adhesive Dentistry, 2*(3), 235-238.

Rosenstiel, S. F., Land, M. F., & Crispin, B. J. (1998). Dental luting agents: A review of the current literature. *Journal of Prosthetic Dentistry, 80*(3), 280-301.

Sarafianou, A., Seimenis, I., & Papadopoulos, T. (2008). Effectiveness of different adhesive primers on the bond strength between an indirect composite resin and a base metal alloy. *The Journal of Prosthetic Dentistry, 99*(5), 377-387.

Schmidlin, P. R., Stawarczyk, B., Wieland, M., Attin, T., Hämmerle, C. H., & Fischer, J. (2010). Effect of different surface pre-treatments and luting materials on shear bond strength to peek. *Dental Materials, 26*(6), 553-559.

Stawarczyk, B., Krawczuk, A., & Ilie, N. (2015). Tensile bond strength of resin composite repair in vitro using different surface preparation conditionings to an aged cad/cam resin nanoceramic. *Clinical Oral Investigations, 19*(2), 299-308.

Su, N., Yue, L., Liao, Y., Liu, W., Zhang, H., Li, X., … & Shen, J. (2015). The effect of various sandblasting conditions on surface changes of dental zirconia and shear bond strength between zirconia core and indirect composite resin. *The Journal of Advanced Prosthodontics, 7*(3), 214-223.

Swift Jr, E., LeValley, B., & Boyer, D. (1992). Evaluation of new methods for composite repair. *Dental Materials, 8*(6), 362-365.

Toh, C., Setcos, J., & Weinstein, A. (1987). Indirect dental laminate veneers - An overview. *Journal of Dentistry, 15*(3), 117-124.

Völkel, T. (2004). *Scientific documentation: multilink automix*. Schaan, Liechtenstein: Ivoclar Vivadent AG.

Wiegand, A., Stucki, L., Hoffmann, R., Attin, T., & Stawarczyk, B. (2015). Repairability of cad/cam high-density pmma-and composite-based polymers. *Clinical Oral Investigations, 19*(8), 2007-2013.

Yavuz, T., Dilber, E., Kara, H. B., Tuncdemir, A. R., & Ozturk, A. N. (2013). Effects of different surface treatments on shear bond strength in two different ceramic systems. *Lasers in Medical Science, 28*(5), 1233-1239.

Yenisey, M., Dede, D. Ö., & Rona, N. (2016). Effect of surface treatments on the bond strength between resin cement and differently sintered zirconium-oxide ceramics. *Journal of Prosthodontic Research, 60*(1), 36-46.

Zhou, L., Qian, Y., Zhu, Y., Liu, H., Gan, K., & Guo, J. (2014). The effect of different surface treatments on the bond strength of peek composite materials. *Dental Materials, 30*(8), e209-e215.

Pertanika J. Sci. & Technol. 27 (4): 1561 - 1574 (2019)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Exploring Potential Neuroprotective Properties of Aqueous *Centella asiatica* Extract in Chronic Stress-induced Rats

Muhamad Zulhusni Abdul Wahab[1], Muhammad Zaim Zainal Abidin[1], Fatin Nadzirah Zakaria[1]*, Mohamad Aris Mohd Moklas[2], Zulkhairi Amom[1], Nina Keterina Hashim[1] and Razif Dasiman[1]

[1]*Faculty of Health Sciences, Universiti Teknologi MARA, 42300 UiTM, Bandar Puncak Alam, Selangor, Malaysia*

[2]*Faculty of Medicine and Health Sciences, Universiti Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia*

ABSTRACT

Centella asiatica is one of the traditional herbs consumed by many communities due to its wide range of applications such as treating Parkinsonism, promoting memory enhancement, and preventing oxidative stress. This study was conducted to investigate the neuroprotective potential of aqueous *C. asiatica* extract (CAE) against neurodegeneration induced by chronic stress. Administration of CAE at three different dosages (200 mg/kg/day, 400 mg/kg/day and 800 mg/kg/day) was conducted for a period of 21 days along with exposure to chronic stress using restrainer and forced swimming regimes. The administration of CAE significantly improved the thickness of dentate gyrus and reduced the amount of neuronal cell death at dentate gyrus and CA3 (p<0.05). Additionally, administration of CAE significantly alleviated the expression c-fos protein (p<0.05). Thus, this study highlighted the neuroprotective effect of CAE against neurodegeneration from chronic stress exposure.

Keywords: Centella asiatica, chronic stress, neuroprotectivity

ARTICLE INFO

Article history:
Received: 15 September 2018
Accepted: 04 June 2019
Published: 21 October 2019

E-mail addresses:
zaczulhusni@gmail.com (Muhamad Zulhusni Abdul Wahab)
zaem87@gmail.com (Muhammad Zaim Zainal Abidin)
fatinn9604@puncakalam.uitm.edu.my (Fatin Nadzirah Zakaria)
aris@upm.edu.my (Mohamad Aris Mohd Moklas)
zulkha2992@puncakalam.uitm.edu.my (Zulkhairi Amom)
ninake0402@puncakalam.uitm.edu.my (Nina Keterina Hashim)
razifdasiman@salam.uitm.edu.my (Razif Dasiman)
* Corresponding author

INTRODUCTION

Stress can be defined as a form of adaptive response towards external demands (Sannino et al., 2016). The adaptive response is primarily mediated by the release of corticosteroids, however, continuous exposure to stress can cause the activation of adaptive response to be severely detrimental (Pinto et al., 2015). Previous studies by McEwen (2012) had shown the relationship

ISSN: 0128-7680
e-ISSN: 2231-8526

between the hippocampal atrophy due to stress and functional deficits in hippocampal-dependent functions.

The central nervous system (CNS) comprises brain and spinal cord, is made up of intricate and differentiated structures. The structures, particularly the limbic structure however is very susceptible and sensitive to chronic stress. Exposure to chronic stress can lead to hippocampal degradation and cause the hippocampus to be more vulnerable to injuries (Conrad, 2008). The hippocampus has a very important role in cognitive and memory processes, thus exposure to chronic stress can cause severe impacts on cognition and memory (Pinto et al., 2015).

Therefore, neuroprotection from continuous exposure of stress is essential to preserve the hippocampal integrity and function. The spontaneous regeneration of damaged tissues however is protracted and particularly impeded in relation to the rate of neuronal degradation. Numerous therapies and commercial drugs are known to promote neuroprotection, however it is usually accompanied with adverse effects. Levodopa is commonly known as the neuroprotective therapeutic agent that delays the progression of Parkinson's disease. However, the efficacy of levodopa is only apparent in long-term application; and prolonged administration of levodopa was shown to promote dyskinesia and in worse case, accelerates the progression of Parkinson's disease (Jankovic & Aquilar, 2008). Natural herbs on the other hand, have potential as alternative remedy to promote neuroprotectivity and known for its versatility and wide range of benefits. The insubstantial success and latent side effects from modern drugs also contribute to shift of interest towards natural herbs as an alternative option. The consumption of traditional herbs has been a daily practice particularly in South East Asia and India for promoting good health, prevention of ailment or even boosting their energy levels (Misra et al., 2008).

Centella asiatica, synonym *Hydrocotyle asiatica* is well known particularly in South East Asia and India region, has variety of names which include 'pegaga', 'gotu kola', 'Brahmi' and Indian pennywort (Mishra, 2015). *C. asiatica* belongs to family Apiaceae and native to countries with tropical and subtropical climate such as Malaysia, India, and China (Halimi, 2011). The pharmacological benefits of *C. asiatica* have been well documented and proven to promote wound healing, neuroprotectivity, memory enhancement as well as reducing oxidative stress (Somboonwong et al., 2012; Bylka et al., 2014; Nasir et al., 2011; Sainath et al., 2011).

C. asiatica is gaining recognition as an alternative option due to its potential neuroprotective properties. A study by Haleagrahara and Ponnusamy (2010) had shown that administration of aqueous *C. asiatica* extract was effective in preventing neuro degenerative changes from Parkinsonism. The administration of aqueous *C. asiatica* extract on *Sprague Dawley* rats induced to Parkinsonism was shown to promote neuroprotective effect and antioxidant level in corpus striatum and hippocampus (Haleagrahara & Ponnusamy, 2010). Previous study by Gray et al. (2015) had shown that primary rat hippocampal neurons treated with water extract of *C. asiatica* promoted neuroprotective activities against amyloid-β toxicity.

However, most of previous studies were focused on neuroprotective potential on Alzheimer's disease and Parkinson's disease, and to date, limited knowledge is available on neuroprotectivity against chronic stress, particularly using natural herbs such as *C. asiatica*. Therefore, considering these issues, this study was conducted to investigate the neuroprotective potential of aqueous *C. asiatica* extract (CAE) on the impact of chronic stress condition. To achieve this aim, a chronic stress model was used, followed by potential neurogenesis and histological assessment on the hippocampus.

MATERIALS AND METHODS

Animals

The study was carried out using male Wistar rats weighing between 200 and 220 g. Rats were housed with two animals per cage and maintained in 12:12 h dark and light cycle, constant temperature of 25±2°C, relative humidity of 40% and allowed free access to certified rodent food and water *ad libitum*. All experiments were performed in accordance to protocols reviewed and approved by Institutional Animal Care and Use Committee (IACUC), Universiti Putra Malaysia (ACUC No: UPM/IACUC/AUP-R078/2018).

Preparation of Extract

Collected *C. asiatica* leaves were authenticated at Atta-ur-Rahman Institute of Natural Product Discovery (AuRIns), Universiti Teknologi MARA (UiTM) and a sample of the plant was deposited (Voucher No: CA-K017). The leaves were dried in hot air oven for three days at 40°C and coarsely ground using mechanical grinder. The preparation of aqueous extract was done using 25 g of coarse plant powder mixed with 250 mL boiling water for 1 h. The mixture was filtered and concentrated into dried powder under vacuum at 50°C.

Experimental Design

Rats were assigned into five different groups (n=8) and supplemented orally for 21 consecutive days as follows. The positive (PC) and negative control (NC) groups were given distilled water. Three groups were given different dosages of CAE respectively as follows, 200 mg/kg/day (CAE 200), 400 mg/kg/day (CAE 400) and 800 mg/kg/day (CAE 800). The dosage was selected based on previous toxicity studies (Oruganti et al., 2010; Chivapat et al., 2011; Deshpande et al., 2015) that showed lethal dose (LD50) of 2000 mg/

kg and no observed adverse effect level (NOAEL) of 1000 mg/kg. Each rat in respective dosage group was fed through oral gavage with the given amount of CAE daily for 21 days. All rat groups with exception to negative control group were left for one hour before being induced to chronic stress for 30 minutes.

The rats were induced into chronic stress using two types of stressor, the restraint and forced swimming regimes alternatively for 21 days (Abidin et al., 2004; Badowska-Szalweska et al., 2010; Sarjan et al., 2017). In restraint stressor regime, the rats were restrained in a cylindrical open-ended plastic restrainer (7 cm in diameter and 22 cm in length) for 30 minutes per day (Sarjan et al., 2017). In forced swimming stressor regime, the rats were forced to swim in a plastic container (50 cm in diameter and 70 cm in height) filled with 200 litres of water at room temperature (25±2°C) for 30 minutes per day (Sarjan et al., 2017). The study was carried out for 21 consecutive days and rats were allowed free access to food and water.

Following the last day of treatment, rats were sacrificed subsequently by decapitation. Brains were quickly removed, cleaned with saline solution and right hemisphere were stored at -80°C for protein determination and left hemisphere were stored in 10% buffered formalin for histological findings.

Histological Study of Hippocampus

Brain tissues were fixed in 10% buffered formalin for seven days. Paraffin blocks were made in embedding bath and sectioned at 5 microns using rotary microtome. The sections were stained with 0.1% cresyl violet. 0.1 g of cresyl violet acetate was dissolved in 100 mL distilled water and 10 drops of glacial acetate acid was added to give pH of 3.5 (Taib, 2014). The stain was filtered prior to usage. The slide was examined using compound research microscope complete with digital image acquisition system (Leica DM1000, Leica Microsystems, Germany) to observe the thickness of dentate gyrus (DG) and neuronal cell death at DG and CA3. The neuronal cell death was scored according to scoring system as shown in Table 1 (Taib, 2014). The DG and CA3 regions of the hippocampus are as shown in Figure 1.

Table 1
The neuronal cell death scoring system and its descriptions

Score	Description
0	No neuronal cell death in field of view or normal
1	A few or sporadically present neuronal cell death in the field of view or normal
2	Group of neuronal cell death in the field of view
3	Group of neuronal cell death in the field of view with minor destruction
4	Massive alteration or all neurons were damaged in the field of view

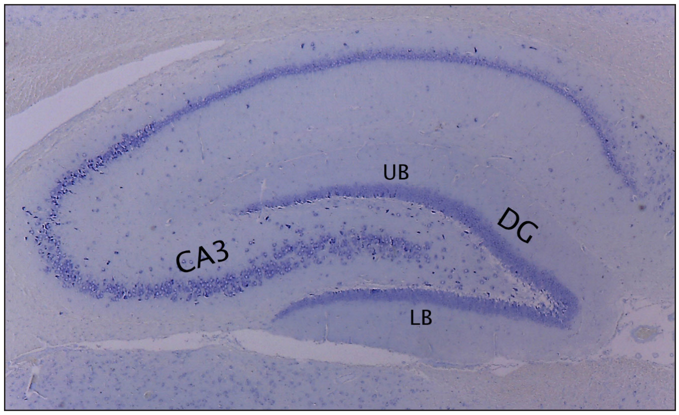


Figure 1. Cresyl violet-stained section of the hippocampus, containing dentate gyrus (DG) and CA3. The DG comprises of upper blade (UB) and lower blade (LB). The section was observed under 5x total magnification.

Total c-fos Protein Level

Sodium dodecyl sulfate-polyacrylamide (SDS) (10%) gels were used according to molecular weight of c-fos protein, 57 kDa. SDS gel electrophoresis were run at 100 V for 40 minutes followed by 150 V at 60 minutes. The separated protein was transferred onto nitrocellulose membrane at 100 V for 60 minutes. The membrane was incubated for 2 h in blocking solution (5% w/v skimmed milk in Tween20) and overnight in primary anti-c-fos antibody (1:1000 dilution, Santa Cruz Biotechnology, Germany) in similar blocking solution. The membrane was then incubated with horseradish peroxidase (HRP)-conjugated secondary antibodies anti-rabbit IgG (1:2000 dilution, Santa Cruz Biotechnology, Germany) for 2 h. Membrane was incubated with chemiluminescent detection reagent (WesternBright^{TM} ECL and Peroxide, Advansta, USA) for 5 minutes and exposed to gel documentation (INFINITY System, Vilber Lourmat, Germany) for 10 minutes. The quantification of c-fos protein was done using ImageJ system protocols (Davarinejad, 2013).

Statistical Analysis

The data were analysed using SPSS software (version 21, IBM SPSS, USA) and significant difference between groups was examined using analysis of variance (ANOVA) followed by individual comparison using Student's *t*-test (two-tailed). All data was expressed as mean ± SD (standard deviation). The values of $p<0.05$ were considered statistically significant.

RESULTS AND DISCUSSION

Histological Study of Hippocampus

The thickness of dentate gyrus (DG) of male Wistar rats was measured and the results were as shown in Figure 2. Group CAE 800 recorded the highest thickness of DG, followed by group CAE 400, group CAE 200, group NC and group PC. The thickness of DG in group

PC was observed to decrease significantly with comparison to group NC (p<0.05). The thickness of DG in group CAE 200, CAE 400 and CAE 800 was significantly higher when compared to group PC (p<0.05). The thickness of the dentate gyrus of group CAE 800 was found to be significantly higher when compared to group CAE 200 (p<0.05).

The neuronal cell death assessment was performed at hippocampal structures of DG and CA3. Any abundance or appearance of neuronal cell death or alteration of neurons observed was noted and the results were as shown in Table 2, Figure 3 and Figure 4. For DG region as shown in Figure 3, Group PC recorded the highest amount of neuronal cell death with scoring of 3 (2.79 ± 0.40), followed by group CAE 200 with neuronal cell death score of 2 (2.24 ± 0.53), group CAE 400 with neuronal cell death score of 2 (1.83 ± 0.84), group CAE 800 with neuronal cell death score of 1 (1.43 ± 0.41) and group NC with neuronal cell death score of 1 (0.95 ± 0.73). The amount of neuronal cell death in group PC and group CAE 200 were significantly higher when compared to group NC (p<0.05). The amount of neuronal cell death in group CAE 400 and CAE 800 were significantly lower when compared to group PC (p<0.05). The amount of neuronal cell death between CAE 400 and CAE 800 however, were not significant when compared to group NC (p>0.05). The amount of neuronal cell death in group CAE 200 was significantly higher when compared to group CAE 800 (p>0.05).

As for CA3 region as shown in Figure 4, Group PC recorded the highest amount of neuronal cell death with scoring of 3 (2.50 ± 0.74), followed by group CAE 200 with neuronal cell death score of 2 (2.14 ± 0.69), group CAE 400 with neuronal cell death score of 2 (1.88 ± 0.83), group CAE 800 with neuronal cell death score of 1 (1.33 ± 1.00) and group NC with neuronal cell death score of 1 (0.86 ± 0.49). The amount of neuronal cell death in group PC and CAE 200 were significantly higher when compared to group NC (p<0.05). The amount of neuronal cell death in group CAE 400 and CAE 800 were significantly lower when compared to group PC (p<0.05). The amount of neuronal cell death between CAE 400 and CAE 800 however, were not significant when compared to group NC (p>0.05).

Table 2
The neuronal cell death score in DG and CA3 region of the hippocampus

Hippocampal Region	NC	PC	CAE 200	CAE 400	CAE 800
DG	1 (0.95±0.73)	3^{a} (2.79±0.40)	2^{a} (2.24±0.53)	2^{b} (1.83±0.84)	$1^{b, c}$ (1.43±0.41)
CA3	1 (0.86±0.90)	3^{a} (2.50±0.74)	2^{a} (2.14±0.69)	2^{b} (1.88±0.83)	1^{b} (1.33±1.00)

Note. Values in brackets were expressed as mean ± SD. ^{a}, p<0.05, compared to NC; ^{b}, p<0.05, compared to PC; ^{c}, p<0.05, compared to CAE 200. NC = negative control, PC = positive control, CAE = *Centella asiatica* extract, 200 = 200 mg/kg/day, 400 = 400 mg/kg/day, 800 = 800 mg/kg/day, DG = dentate gyrus.

The brain contains a specific limbic structure known as the hippocampus that plays an essential role in memory and cognitive processes (Fanselow & Dong, 2010). Specific hippocampal structures known as DG and CA3 function as the centre for memory and learning processes (Gao et al., 2007). However, DG and CA3 are very susceptible to neurodegeneration and severe condition of neurodegeneration could lead to functional

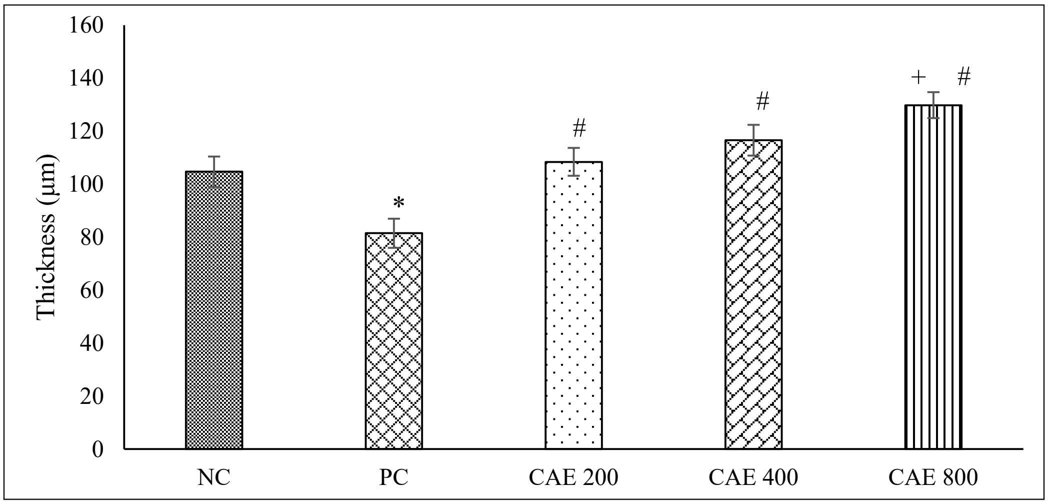


Figure 2. The thickness of dentate gyrus. Data were expressed as mean ± SD. NC = negative control, PC = positive control, CAE = *Centella asiatica* extract, 200 = 200 mg/kg/day, 400 = 400 mg/kg/day, 800 = 800 mg/kg/day. (* indicates p<0.05 vs group NC; # indicates p<0.05 vs group PC; + indicates p<0.05 vs group CAE 200).

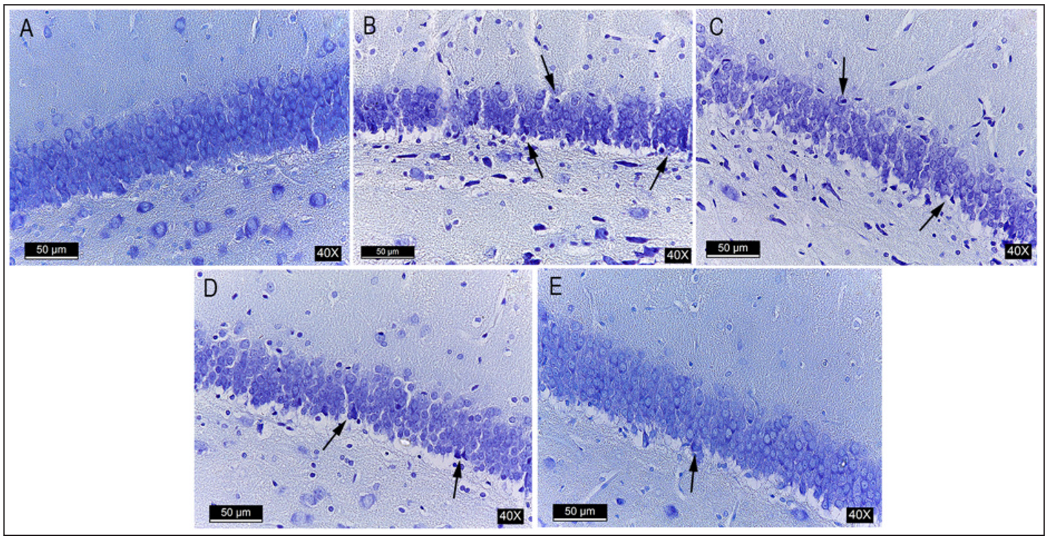


Figure 3. Cresyl violet-stained sections of dentate gyrus. Cells were observed under 40x total magnification. Arrows indicate neuronal cell death. A = negative control, B = positive control, C = CAE 200 (200 mg/kg/day), D = CAE 400 (400 mg/kg/day), E = CAE 800 (800 mg/kg/day).

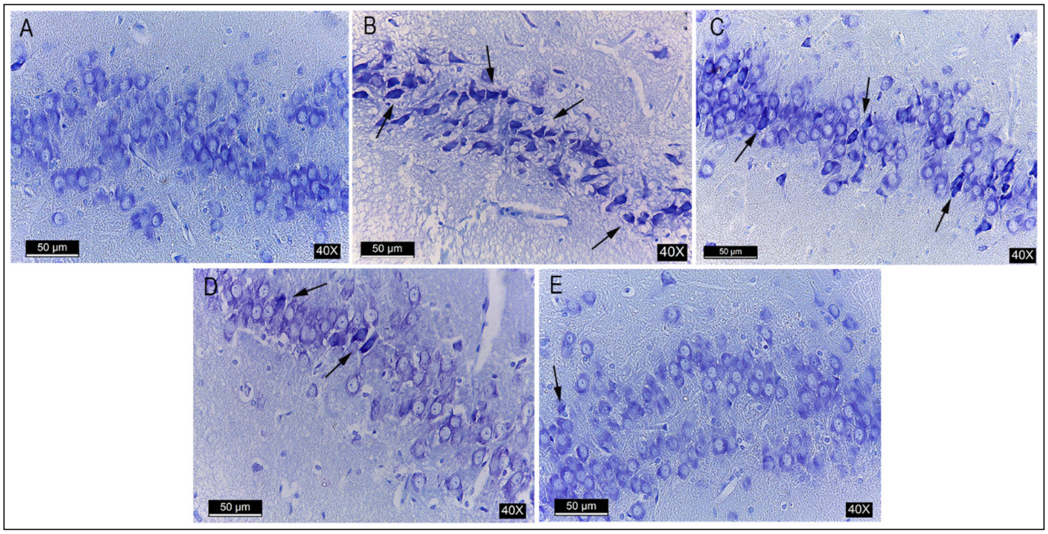


Figure 4. Cresyl violet-stained sections of CA3. Cells were observed under 40x total magnification. Arrows indicate neuronal cell death. A = negative control, B = positive control, C = CAE 200 (200 mg/kg/day), D = CAE 400 (400 mg/kg/day), E = CAE 800 (800 mg/kg/day).

impairment of the hippocampus (Chohan et al., 2009). Continuous exposure to stress can cause the degeneration of hippocampal structures and disrupts the hippocampal functions (Sousa & Almeda, 2012). Despite so, the remission of neurodegeneration of the hippocampus can be achieved through stimulation of neurogenesis or neuroprotection. Therefore, the neuroprotective potential of CAE was studied by assessing its potential shielding effect and neurogenesis upon exposure to chronic stress.

The thickness of DG was observed particularly at the upper blade (UB) and the lower blade (LB), as any fluctuation on its thickness will reflect on the rate of the neurogenesis (Li et al., 2008). Any elevation of neuronal formation activities or neurogenesis can be represented by increasing thickness of the DG and vice versa (Lie et al., 2004). The neuroprotective potential of CAE was further assessed through observation on the amount of neuronal cell death in DG and CA3. The DG and CA3 structures were observed due to their roles in memory and cognitive processes, as well as their susceptibility to neurodegeneration (Evstratova & Tóth, 2014). The observation on the amount of neuronal cell death would translate into potential rate of neuroproliferation at observed hippocampal structures. Surfeited group of neuronal cell death or conspicuous destruction of the prescribed structures are usually accompanied by reduction in size of the hippocampal structures (Langmeier et al., 2003).

The observation on DG as shown in Figure 2 showed a significant decrease in thickness particularly in group PC when compared to group NC (p<0.05). The observation on group PC expressed the highest neuronal cell death score in both DG and CA3 as shown in Table 2. The observations postulated the neurodegenerative condition of the hippocampus and agrees with the study by Hashemet al. (2010). The results indicated that the chronic stress model was successful in exhibiting neurodegenerative condition on hippocampal structures, by which there was apparent neuronal loss after 21 days of exposure to chronic stress.

The administration of CAE in the different dosages (CAE 200, CAE 400 and CAE 800) showed a significant increase in thickness of DG with comparison to group PC as shown in Figure 2 (p<0.05). The evaluation was reflected with the observations in group CAE 200, CAE 400 and CAE 800 that showed a significant decrease of neuronal cell death abundance, particularly DG when compared to group PC (p<0.05). The results were supported by similar observations in previous studies (Hashem et al., 2010; Abdallah et al., 2010) which reflected the evidence of neurogenesis in hippocampal structures by increased neuroproliferation or reduced amount of neuronal cell death as well as improvement on DG thickness. The results showed that the administration of CAE promoted neurogenesis activities in hippocampus upon exposure to chronic stress, indicating a potential neuroprotective effect of the extract.

The administration of CAE in group CAE 200, CAE 400 and CAE 800 respectively showed improving proliferative pattern on DG and CA3, as reflected with lower neuronal cell death score as shown in Table 2. Despite so, the administration of CAE did not show any significant improvement on the neuroproliferation in DG and CA3 when compared to group NC (p>0.05). The observation suggested that the administration of CAE showed a recovery pattern on DG and CA3 upon exposure to chronic stress that corresponded to those in group NC. Similar observations were found in previous study by Madhyasta et al. (2007) which demonstrated neuroprotective effect of fresh *C. asiatica* extract in prenatal stress-induced rats. The administration of fresh *C. asiatica* extract was shown to minimise neuronal loss from prenatal stress and enhanced the neuronal proliferative activities in the hippocampus. Another study by Perederiy and Westbrook (2013) demonstrated that the neuronal regeneration from eventuating injuries on the hippocampal structures was capable of retaining normal hippocampal functions as well achieving structural neuroplasticity. The results supported that continuous supplementation of CAE could reduce or minimise the detrimental consequences from daily chronic stress exposure in human.

Total c-fos Protein Level

The obtained protein bands of the total c-fos protein were quantified and the results were as shown in Figure 5. The percentage changes of total c-fos protein among group PC, CAE 200, CAE 400 and CAE 800 were expressed as percentage over group NC. Group CAE 800 recorded the highest level of total c-fos protein with 175.63 ± 43.60 % changes over group NC, followed by group CAE 400 with 107.25 ± 34.63 % changes over group NC, and group CAE 200 with 75.26 ± 20.43 % changes over group NC. Group PC recorded a reduction of 6.29 ± 19.76 % changes over group NC, however, no significant difference was recorded between group PC and group NC (*p*>0.05). The expression of total c-fos protein group CAE 200, CAE 400 and CAE 800 showed significant differences in comparison to group NC and PC (*p*<0.05). The expression of total c-fos protein in group CAE 800 is significantly higher with comparison to group CAE 200 (*p*<0.05).

The c-fos is commonly used as a marker in determining the neuronal activity and behavioural response towards acute stimuli exposure (Velazquez et al., 2015a). Another study by Velazquez et al. (2015b) showed that c-fos played a significant role in neurogenesis. Neurogenesis is the formation of new neurons that occurs in constant manner particularly in the hippocampus (Velazquez et al., 2015b). c-fos is made up of two main domains: (1) basic domain (BD) which plays an important role in DNA-binding, and (2) leucine zipper domain (LZ) which plays an important role in heterodimerisation of c-fos with other leucine zipper-containing proteins such as *c-jun* (Curran & Morgan, 1995; Angel & Karim, 1991). The heterodimerisation of c-fos leads to the formation of AP-1 (Activator Protein-1) complex (Vivar & Van Praag, 2013). The AP-1 complex has a significant role in regulating the cellular processes such as differentiation and proliferation (Verma & Graham, 1987).

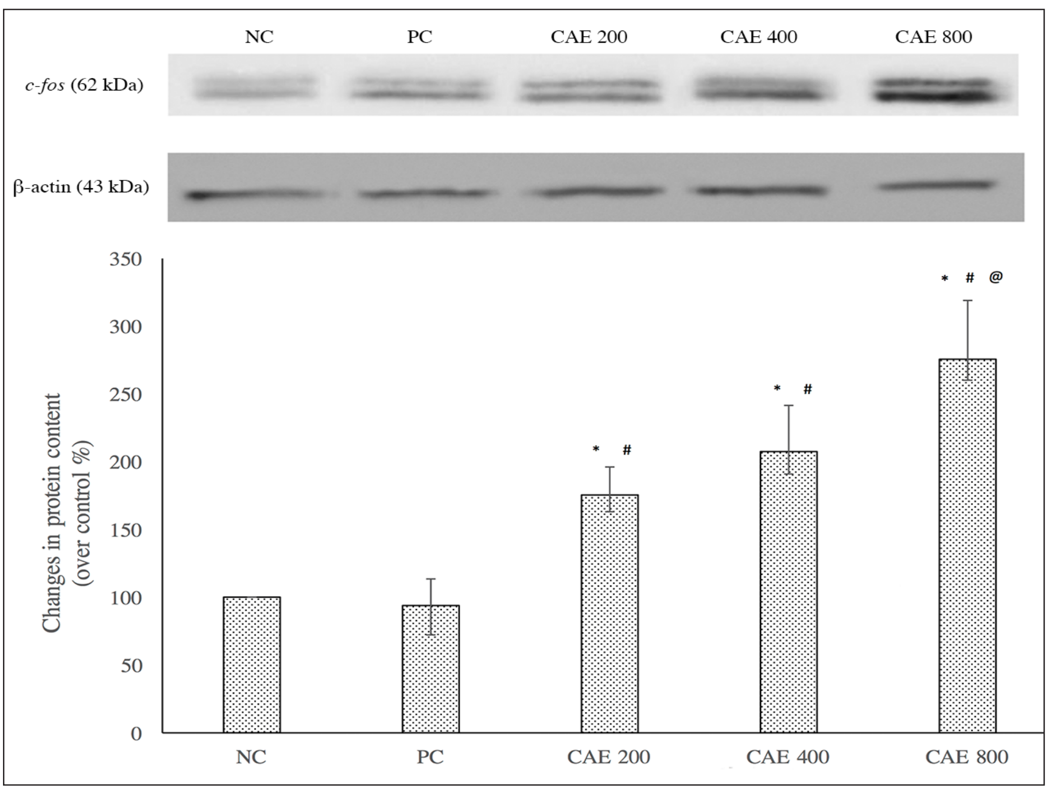


Figure 5. Relative expression of c-fos protein. Data were expressed as percentage change (mean ± SD). Histograms with symbols were significantly different at *p*<0.05 (*, *p*<0.05, compared to NC; #, *p*<0.05, compared to PC; @, *p*<0.05, compared to CAE 200). NC = negative control, PC = positive control, CAE = *Centella asiatica* extract, 200 = 200 mg/kg/day, 400 = 400 mg/kg/day, 800 = 800 mg/kg/day.

Another study by Caubet (1989) demonstrated that the expression of c-fos at various stages of neurodevelopment entailed to the maturation of brain structures. Therefore, elevation of the expression of c-fos protein level may signify an increasing activity of neurogenesis in the hippocampus.

The significant elevation of c-fos expression denotes that CAE promotes the elevation of hippocampal neurogenesis activities in male Wistar rats exposed to chronic stress. The results are supported by previous study by Clark et al. (2010) which suggested that the elevation of neurogenesis activities was in concomitant to an increase in c-fos expression observed in the hippocampus. Clark et al. (2010) concluded that the rate of neurogenesis in the hippocampus was reflected with c-fos expression, by which higher neuronal activities and neurogenesis were parallel to upregulation of hippocampal c-fos. Interestingly, the increasing dosage of CAE demonstrated an increasing pattern of c-fos protein expression. The relative expression of c-fos in group CAE 800 showed the highest expression of c-fos. Therefore, the results denoted that the administration of CAE at high dosage promotes the optimum elevation of neurogenesis activities and was reflected on histological findings on the hippocampus.

CONCLUSION

The administration of CAE at three different dosages exhibited significant neurogenesis activities through the apparent thickening of dentate gyrus and improved neuroproliferation in the hippocampus. The apparent improvement in neurogenesis activities was reflected by significant elevation pattern of c-fos protein expression in groups administered with CAE.

The results of this study have demonstrated the neuroprotective effect of continuous administration of CAE against degeneration of hippocampus structures upon exposure to chronic stress. Therefore, continuous supplementation of *Centella* asiatica could reduce or minimise the detrimental consequences from daily chronic stress exposure in human. Thus, the neuroprotectivity potential of *Centella asiatica* extract shows a promising way as an alternative option to treat chronic stress. The dosage of 800 mg/kg of aqueous *Centella asiatica* extract demonstrated as the optimum dosage for neuroprotective properties in ameliorating the effects of chronic stress.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the support given by Ministry of Agriculture (MoA), Malaysia (Project Number: 100-MOA/RMI/16/6/2); as well as Universiti Teknologi MARA for Geran Inisiatif Penyeliaan (GIP) (Project Number:600-IRMI/MyRA 5/3/GIP) and Geran Penyelidikan LESTARI (Project Number: 600-IRMI/MyRA 5/3/LESTARI). The authors are also thankful to Universiti Teknologi MARA (UiTM) and Universiti Putra Malaysia (UPM) for the support and facilities used during this study.

REFERENCES

Abdallah, N. B. A., Slomianka, L., Vyssotski, A., & Lipp, H. P. (2010). Early age-related changes in adult hippocampal neurogenesis in C57 mice. *Neurobiology of Aging*, *31*(1), 151-161.

Abidin, I., Yargiçoglu, P., Agar, A., Gümüslü, S., Aydin, S., Oztürk, O., & Sahin, E. (2004). The effect of chronic restraint stress on spatial learning and memory: relation to oxidant stress. *International Journal of Neuroscience*, *114*(5), 683-699.

Angel, P., & Karim, M. (1991). The role of *Jun*, *Fos* and the AP-1 complex in cell-proliferation and transformation. *Biochimica et Biophysica Acta*, *1072*(2-3), 129-157.

Badowska-Szalweska, E., Spodnik, E., Klejbor, I., & Morys, J. (2010). Effects of chronic forced swim stress on hippocampal brain-derived neutrophic factor (BDNF) and its receptor (TrkB) immunoreactive cells in juvenile and aged rats. *Acta Neurobiologiae Experimentalis*, *70*(4), 370-381.

Bylka, W., Znajdek-Awiżeń, P., Studzińska-Sroka, E., Dańczak-Pazdrowska, A., & Brzezińska, M. (2014). *Centella asiatica* in dermatology: An overview. *Phytotherapy Research*, *28*(8), 1117-1124.

Caubet, J. (1989). *c-fos* proto-oncogene expression in the nervous system during mouse development. *Molecular and Cellular Biology*, *9*(5), 2269-2272.

Chivapat, S., Chavalittumrong, P., & Tantisira, M. H. (2011). Acute and sub-chronic toxicity studies of a standardized extract of *Centella asiatica* ECa 233. *Thai Journal of Pharmaceutical Sciences*, *35*(2011), 55-64.

Chohan, M. O., Li, B., Blanchard, J., Tung, Y. C., Heany, A. T., Rabe, A., … & Grundke-Iqbal, I. (2009). Enhancement of dentate gyrus neurogenesis, dendritic and synaptic plasticity and memory by a neurotrophic peptide. *Neurobiology of Aging*, *32*(8), 1420-1434.

Clark, P., Kohman, R., Miller, D., Bhattacharaya, T., Haferkamp, E., & Rhodes, J. (2010). Adult hippocampal neurogenesis and *c-Fos* induction during escalation of voluntary wheel running in C57BL/6J mice. *Behavioural Brain Research*, *213*(2), 246-252.

Conrad, C. D. (2008). Chronic stress-induced hippocampal vulnerability: the glucocorticoid vulnerability hypothesis. *Reviews in the Neurosciences*, *19*(6), 395-411.

Curran, T., & Morgan, J. (1995). Fos: an immediate-early transcription factor in neurons. *Journal of Neurobiology, 26*(3), 403-412.

Davarinejad, H. (2013). *Quantification of Western Blots with ImageJ*. Retrieved January 17, 2018, from http://www.yorku.ca/yisheng/Internal/Protocols/ImageJ.pdf

Deshpande, P. O., Mohan, V., & Thakurdesai, P. (2015). Preclinical Safety Assessment of standardized extract of *Centella asiatica* (L.) Urban leaves. *Toxicology International*, *22*(1), 10-20.

Evstratova, A., & Tóth, K. (2014). Information processing and synaptic plasticity at hippocampal mossy fiber terminals. *Frontiers in Cellular Neuroscience*, *8*(28), 1-12.

Fanselow, M., & Dong, H. W. (2010). Are the dorsal and ventral hippocampus functionally distinct structures? *Neuron*, *65*(1), 7-19.

Gao, X., Arlotta, P., Macklis, J., & Chen, J. (2007). Conditional knock-out of β-catenin in postnatal-born dentate gyrus granule neurons results in dendritic malformation. *Journal of Neuroscience*, *27*(52), 14317-14325.

Gray, N. E., Sampath, H., Zweig, J. A., Quinn, J. F., & Soumyanath, A. (2015). *Centella asiatica* attenuates Amyloid-β-Induced oxidative stress and mitochondrial dysfunction. *Journal of Alzheimer's Disease*, *45*(3), 933-946.

Haleagrahara, N., & Ponnusamy, K. (2010). Neuroprotective effect of *Centella asiatica* extract (CAE) on experimentally induced parkinsonism in aged Sprague-Dawley rats. *Journal of Toxicological Sciences*, *35*(1), 41-47.

Halimi, M. S. (2011). Identification of agronomic traits of *Centella asiatica* (L.) Urban naturally grown at regions with different altitude. *Jurnal Natur Indonesia*, *13*(65), 232-236.

Hashem, H., Elmasry, S., & Eladl, M. (2010). Dentate gyrus in aged male albino rats (histological and tau-immunohistochemical study). *Egyptian Journal of Histology*, *33*(4), 659-670.

Jankovic, J., & Aquillar, L. G. (2008). Current approaches to the treatment of Parkinson's disease. *Neuropsychiatric Disease and Treatment*, *4*(4), 743-757.

Langmeier, M., Folbergrová, J., Haugvicová, R., Pokorný, J., & Mares, P. (2003). Neuronal cell death in hippocampus induced by homocysteic acid in immature rats. *Epilepsia*, *44*(3), 299-304.

Li, B., Yamamori, H., Tatebayashi, Y., Shafit-Zagardo, B., Tanimukai, H., Chen, S., … & Grundke-Iqbal, I. (2008). Failure of neuronal maturation in Alzheimer disease dentate gyrus. *Journal of Neuropathology and Experimental Neurology*, *67*(1), 78-84.

Lie, D., Song, H., Colamarino, S., Ming, G., & Gage, F. (2004). Neurogenesis in the adult brain: New strategies for central nervous system diseases. *Annual Review of Pharmacolgy and Toxicology*, *44*, 399-421.

Madhyasta, S., Somayaji, S. N., Bairy, K. L., & Madhyasta, P. (2007). Neuroprotective effect of *Centella asiatica* leaf extract treatment on cognition and hippocampal morphology against prenatal stress. *Thai Journal of Physiological Sciences*, *20*(2), 79-88.

McEwen, B. (2012). The ever-changing brain: cellular and molecular mechanisms for the effects of stressful experiences. *Developmental Neurobiology*, *72*(6), 872-90.

Mishra, S. (2015). Ethnopharmacological importance of *Centella asiatica* with special reference to neuroprotective activity. *Asian Journal of Pharmacology and Toxicology*, *3*(10), 49-53.

Misra, S., Maikhuri, R., Kala, C., Rao, K., & Saxena, K. G. (2008). Wild leafy vegetables: A study of their subsistence dietic support to the inhabitants of Nanda. *Indian Journal of Ethnobiology*, *4*(15), 15-23.

Nasir, M. N., Habsah, M., Zamzuri, I., Rammes, G., Hasnan, J., & Abdullah, J. (2011). Effects of asiatic acid on passive and active avoidance task in male Spraque-Dawley rats. *Journal of Ethnopharmacology*, *134*(2), 203-209.

Oruganti, M., Roy, B. K., Singh, K. K., Prasad, R., & Kumar, S. (2010). Safety assessment of *Centella asiatica* in albino rats. *Pharmacognosy Journal*, *2*(16), 5-13.

Perederiy, J., & Westbrook, G. (2013). Structural plasticity in the dentate gyrus- revisiting a classic injury model. *Front Neural Circuit*, *7*(17), 1-11.

Pinto, V., Costa, J., Morgado, P., Mota, C., Miranda, A., Bravo, F. V., … & Sousa, N. (2015). Differential impact of chronic stress along the hippocampal dorsal-ventral axis. *Brain Structure and Function*, *220*(2), 1205-1212.

Sainath, S. B., Meena, R., Supriya, C., Reddy, K. P., & Reddy, P. S. (2011). Protective role of *Centella asiatica* on lead-induced oxidative stress and suppressed reproductive health in male rats. *Environmental Toxicology and Pharmacology*, *32*(2), 146-154.

Sannino, G., Pasqualini, L., Ricciadelli, E., Montilla, P., Soverchia, L., Ruggeri, B., … & Hardiman, G. (2016). Acute stress enhances the expression of neuroprotection- and neurogenesis-associated genes in the hippocampus of a mouse restraint model. *Oncotarget*, *7*(8), 8455-8465.

Sarjan, H. N., Divyashree, S., & Yajurvedi, H. N. (2017). The protective effect of the Vacha rhizome extract on chronic stress-induced immunodeficiency in rat. *Pharmaceutical Biology*, *55*(1), 1358-1367.

Somboonwong, J., Kankaisre, M., Tantisira, B., & Tantisira, M. H. (2012). Wound healing activities of different extracts of *Centella asiatica* in incision and burn wound models: an experimental animal study. *BMC Complementary and Alternative Medicine*, *12*(1), 103-109.

Sousa, N., & Almeda, O. (2012). Disconnection and reconnection: the morphological basis of (mal)adaptation to stress. *Trends in Neurosciences*, *35*(12), 742-751.

Taib, C. N. M. (2014). *Structural Changes of Hippocampi and Expression of HSP70, c-fos and DREAM Protein in the Spinal Cord after Acute Thermal Stimulus of Early Stress-Experienced Rats* (PhD thesis). Universiti Putra Malaysia, Malaysia.

Velazquez, F., Caputto, B., & Boussin, F. (2015). *c-Fos* importance for brain development. *Aging*, *7*(12), 1028-1029.

Velazquez, F., Prucca, C., Etienne, O., D'Astolfo, D., Silvestre, D., Boussin, F., & Caputto, B. (2015). Brain development is impaired in *c-fos* −/− mice. *Oncotarget*, *6*(19), 16883-16901.

Verma, I., & Graham, W. (1987). The fos oncogene. *Advances in Cancer Research*, *49*, 29-52.

Vivar, C., & Van Praag, H. (2013). Functional circuits of new neurons in the dentate gyrus. *Frontiers in Neural Circuits*, *7*(15), 1-13.

Pertanika J. Sci. & Technol. 27 (4): 1575 - 1588 (2019)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Forensic Parameter and Paternity Testing of Tpa-25 Element in Kelantan-Malay and Jawa-Malay

Nurul Syafika Muslimin[1], Nur Hilwani Ismail[1] and Farida Zuraina Mohd Yusof[1,2]*

[1]School of Biology, Faculty of Applied Sciences, Universiti Teknologi MARA, 40450 UiTM, Shah Alam, Selangor, Malaysia
[2]Integrative Pharmacogenomics Institute (iPROMISE), 42300 Bandar Puncak Alam, Shah Alam, Selangor, Malaysia

ABSTRACT

Malays consist of multi sub-ethnic group believed to have different ancestral origins based on their migrations centuries ago. The DNA profiling for every individual in Malaysia is not recorded, making Malaysia lacking in genetic data of its own citizens. This research aimed to study the geographic-ancestry origin of two Malay sub-ethnic population; Kelantan-Malay and Jawa-Malay by looking into the variation of TPA-25 insertion in each population. It specifically studied on several areas of Peninsular Malaysia in the region of Kelantan, Selangor and Johor as the representative of main areas with high percentage of Kelantan-Malay and Jawa-Malay populations. All the data were obtained from an application of TPA-PCR method, forensic parameter (F-statistic) and survey questionnaire that polled genetically on their ancestry origin in each sub-ethnic population. The research showed that population with high percentage of heterozygous allele (Tt) of TPA-25 insertion was likely to have high possibility of genetic drift occurrence. Jawa-Malay showed the highest percentage of heterozygous allele (Tt) with approximately 48% of the population. The FIS value of Kelantan-Malay and Jawa-Malay populations were recorded positive with the values of 0.678 and 0.366 respectively.

ARTICLE INFO

Article history:
Received: 10 November 2018
Accepted: 4 June 2019
Published: 21 October 2019

E-mail addresses:
pyka07.ns@gmail.com (Nurul Syafika Muslimin)
hilwani.ismail@gmail.com (Nur Hilwani Ismail)
fzuraina@salam.uitm.edu.my (Farida Zuraina Mohd Yusof)
*Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

Moreover, the FIT value recorded was 0.535 which suggested that these two populations were deficits of heterozygotes.

Keywords: F-statistic, genetic drift, heterozygosity, TPA-25 element

---

INTRODUCTION

Malays (Melayu) belongs to an ethnic group who speak Malayo - Polynesian language which is a member of the Austronesian family (Bellwood, 2007; Omar, 2004). They dominantly inhabit the Malay Peninsula, the east coast of Sumatra and the coast of Borneo (Bellwood, 2007). In Peninsular Malaysia, the Malays consist of multi sub-ethnic group believed to have different ancestral origins based on their migrations centuries ago (Wheatley, 1961). The Malay Peninsula was once a very strategic port and trading centre, connecting Indochina and the Indonesian archipelago (Jacq-Hergoualc'h, 2018). However, migrating populations from surrounding areas have further confounded the investigation of the origin of Malays. Recently, there has been deep interest in using Alu elements for application to forensic casework in studies of human population genetic structure and inference of individual geographic origin (Asari et al., 2012; Bamshad et al., 2003; Ray et al., 2005). Alu elements are transposable elements which reach over one million copies in human genome (Ade et al., 2013; Cordaux et al., 2007). These elements are approximately 300 base pairs sequence long which have expanded in human genome for more than 60 million years (Batzer & Deininger, 2002). Alu elements are characterized by their ability to "copy and paste" via a mechanism to produce new copies by using RNA transcript to be reverse trancripted as cDNA and the duplicate is inserted at a new genomic location (Batzer & Deininger, 2002).

These Alu elements are expanded in human genome at a substantial rate (Cordaux et al., 2006) resulting in multiple copies in human genome, making them as the most successful mobile elements (Batzer & Deininger, 2002; Lander et al., 2001) and represented as an important source of human genomic variation (Batzer & Deininger, 2002). TPA-25 sequence is an Alu element within tissue plasminogen activator gene, a dimorphic gene not represented in all individuals (Batzer et al., 1991). This marker has been used to study various population in central Argentina Patagonia, North Africa (Tunisian) and South India (Ennafaa et al., 2011; Lotfi et al., 2011; Moncer et al., 2011; Parolin et al., 2017; Veerraju et al., 2001). TPA-25 suitable for both ancestry and admixture analysis because of the present of importance source of information regarding on genetic diversity, both from the current diversity, and throughout the process of human evolution (Cherni et al., 2011; Rishishwar et al., 2015).

These criteria suit the sequence as a better gene marker for screening the genetic structure of human population. This study aimed to determine the variation of TPA-25 insertion in each population of two sub-ethnic groups, Kelantan-Malay and Jawa-Malay, the two groups that were categorized as the one of the biggest ethnic groups in Peninsular Malaysia by using TPA-PCR-based method, to calculate the allele and genotype frequencies of each human DNA sample by using Hardy-Weinberg equation and to determine the geographic-ancestry origin of each population by an application of forensic parameters and parental lineage investigation.

MATERIALS AND METHODS

Sample Collection

Total human genome DNA samples in the form of buccal cell were collected on the FTA Mini Cards (WHATMAN, Germany). Fifty-one samples were from unrelated Kelantan-Malay population and another thirty-three samples from unrelated Jawa-Malay population. Figure 1 shows the geographical map of populations used in this study. The samples were limited within Peninsular Malaysia and collected from different regions. Consideration of looking into their skin colour and appearance was excluded. The ethical approval on sample collection was obtained from the Research and Ethics Committee of Universiti Teknologi MARA (UiTM) [Ref no: 600-RMI (5/1/6/01)].

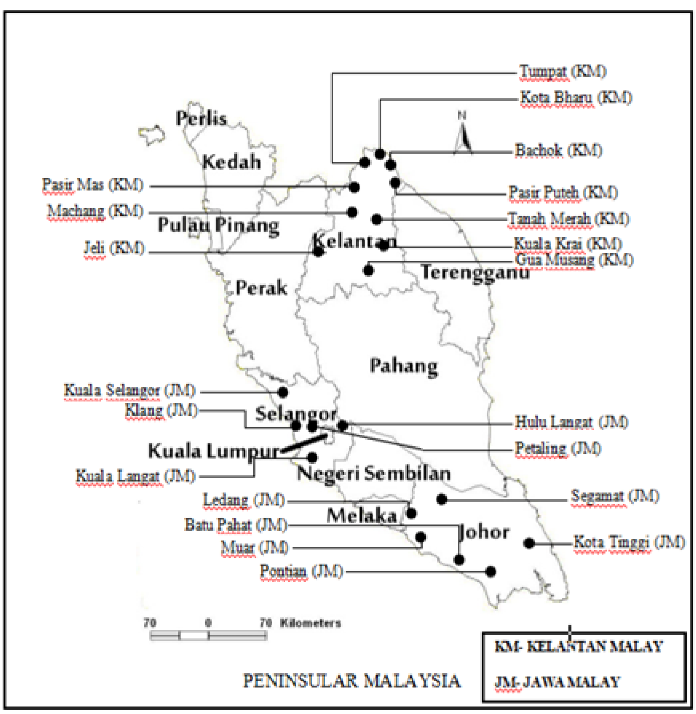


Figure 1. Geographic location of the Malay sub-ethnic groups in Peninsular Malaysia used in this study.

PCR Amplification

Oligonucleotide primers for TPA-25 element PCR were TPA-forward 5'GTAAGAGTTCGTAACAGGACAGCT3' and TPA-reverse 3' CCCCACCCTAGAGAACTTCTCTTT5' with size product of ~400 bp (TPA-25 insertion) and ~100 bp (TPA-25 deletion). PCR conditions were optimized for each assay regarding annealing temperature and oligonucleotide primers.

Amplification of each Alu Polymorphism (TPA-25 sequence) was performed in 22 µl reaction using 2 mm FTA card disc, 10x Optimized DyNAzymes buffer, 20 mM dNTPs (Solis BioDyne), 10 mM oligonucleotide primer (AITBIOTECH PTE LTD, Singapore), DyNazyme II DNA polymerase and double distilled water. Each reaction contained 49 µl of PCR master mix and 1 FTA disc of DNA template. The samples were then subjected to an initial denaturation of 13 minutes at 95 °C followed by 35 amplification cycles of denaturation at 94 °C for 15 seconds, annealing temperature for 30 seconds and followed by extension at 72 °C for 30 seconds. After the final extension at 72 °C for 10 minutes, the samples were kept at 4 °C.

Statistical Analysis

Allele frequency of each samples was calculated by using Hardy-Weinberg equation (Eq.1). The data was represented as the heterozygosity of each Malays sub-ethnic population; Kelantan-Malay and Jawa-Malay.

$$p + q = 1$$
$$(p+q)^2 = p^2 + 2pq + q^2 \qquad [1]$$

The Fixation index, FST value was measured through F-statistics (Roewer et al., 2000). Three levels fixation indices that reflect the biological of organization, individual, subpopulation and the total population (i) inbreeding coefficient (FIS), (ii) fixation index (FST) and (iii) overall inbreeding coefficient (FIT) (Saltkin, 1995).

The reduction in heterozygosity of an individual due to non-random mating within its subpopulation was estimated by following formula of FIS :

$$\text{FIS} = (\text{HS} - \text{HI}) / \text{HS} \qquad [2]$$

It shows the degree to which heterozygosity is reduced below the expectation. The value of FIS ranges between -1 and +1. Negative FIS values indicate heterozygote excess (outbreeding) and positive values indicate heterozygote deficiency (inbreeding) compared with Hardy-Weinberg Equilibrium expectations.

The reduction in heterozygosity of an individual relative to the total population was estimated by following formula of FIT :

$$\text{FIT} = (\text{HT} - \text{HI}) / \text{HT} \qquad [3]$$

The degree of the genetic differentiation between sub-populations was estimated by following formula of FST :

FST = (HT - HS) / HT [4]

FST value is always positive, ranges between 0 (no subdivision, random mating occurring, no genetic divergence within the population) and 1 (complete isolation with extreme subdivision).

RESULTS AND DISCUSSION

The allele frequencies of TPA-25 sequence in each sample for both Malays sub-ethnic populations were determined based on the existence of TPA-25 insertion (~400bp) and TPA-25 deletion (~100bp) which is shown in Figure 2 and Figure 3. Meanwhile, the size estimation of the insertion and deletion are shown in Table 1 and Table 2. After collecting the genotyping data, Table 3 and Table 4 were created to show the distribution of three different types of group which are (i)group of homozygous insertion (+ve), (ii) homozygous deletion (-ve), and (iii) group of heterozygous alleles (+ve/-ve).

Figure 4 shows that from the total of 33 individuals sampled from Jawa-Malay population, 27.27% of the population possessed TPA-25 insertion, 12.12% possessed TPA-25 deletion and the remaining of the population (48.48%) possessed heterozygous allele of TPA-25 element. Meanwhile, from 51 individuals of Kelantan-Malay population, about 17.64% of the population possessed TPA-25 insertion, 37.25% possessed TPA-25 deletion and the remaining of the population (29.41%) possessed heterozygous allele of TPA-25 element.

It shows the heterozygous group in Jawa-Malay population is higher than heterozygous group in Kelantan-Malay. This indicates the high possibility of genetic drift in the Jawa-Malay population as may be due to the result of continuous network of migrations, invasions and admixture of people from different origins (Cherni et al., 2011; Tishkoff & Kidd, 2004). These practices and the ancient movements were active in hundreds of years prior while many different ethnics from neighbouring nations came and abode in Tanah Melayu (Hugo, 1993; Shamsul, 2001).

Eventually, the activities then contributed to the existence of different Malays sub-ethnic population resulting from the intermarriage between the Proto-Malay and other ethnic groups. Different ethnic population also could become unified under singular cultural, religion practices, and language family by simply accepting, adoption and adaptation process (Keita, 2010). Meanwhile, the low heterogeneous gene in a population reflecting their reproductive isolation which might be due to religious and cultural differences seemingly to have provide an obstacle to their intermixing (Cherni et al., 2011; Ennafaa et al., 2006).

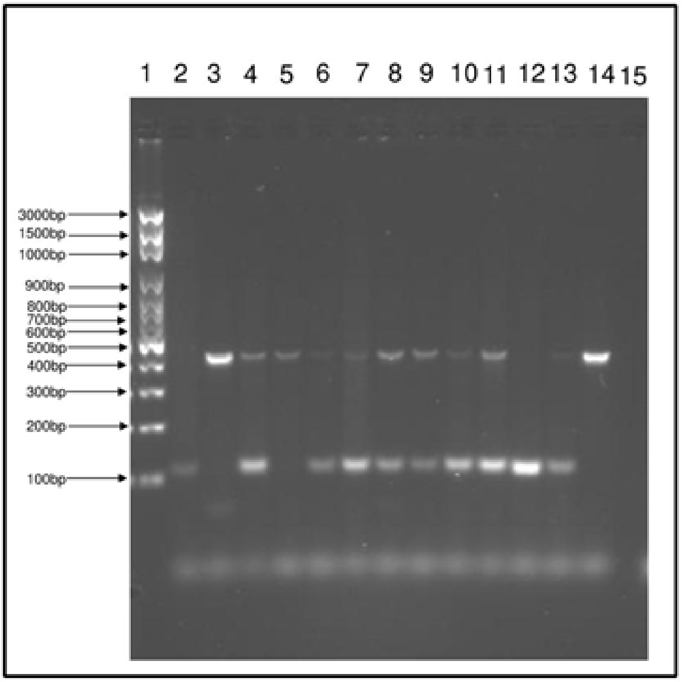


Figure 2. A total of 13 samples of PCR products for TPA-25 sequence of Jawa-Malay population on 2% (w/v) Agarose gel.

Table 1

Size estimation of TPA-25 element of 13 samples of Jawa-Malay population

Lane No.	Sample	TPA-25 Insertion	TPA-25 Deletion
1	100 bp DNA ladder		
2	PCR product of sample 43J	No amplification	114bp
3	PCR product of sample 44J	478bp	69bp
4	PCR product of sample 45J	436bp	114bp
5	PCR product of sample 49J	478bp	No amplification
6	PCR product of sample 50J	478bp	114bp
7	PCR product of sample 52J	478bp	114bp
8	PCR product of sample 53J	478bp	114bp
9	PCR product of sample 54J	478bp	114bp
10	PCR product of sample 55J	524bp	204bp
11	PCR product of sample 56J	478bp	186bp
12	PCR product of sample 57J	No amplification	169bp
13	PCR product of sample 58J	No amplification	186bp
14	PCR product of sample 59J	478bp	No amplification
15	Negative control	No amplification	No amplification

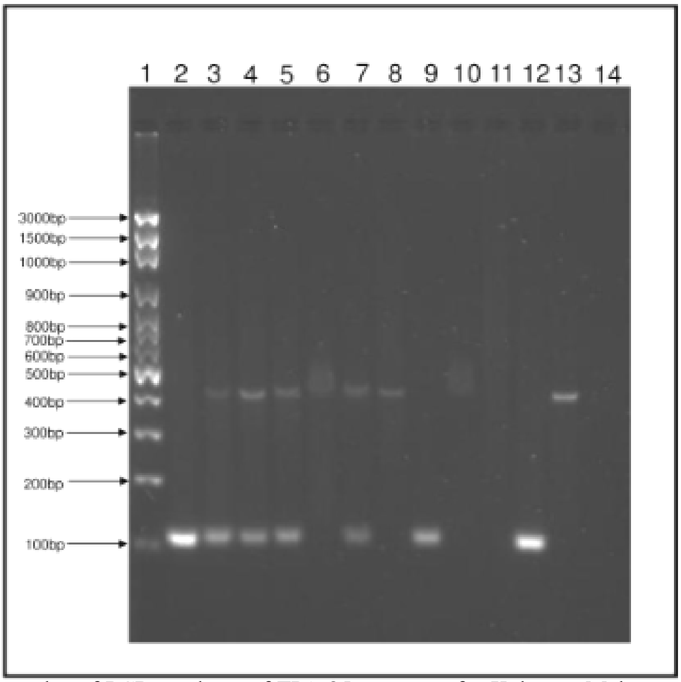


Figure 3. A total 12 samples of PCR products of TPA-25 sequence for Kelantan-Malay population on 2% (w/v) Agarose gel

Table 2

Size estimation of TPA-25 element of 14 samples of Kelantan-Malay population

Lane No.	Sample	TPA-25 Insertion	TPA-25 Deletion
1	100 bp DNA ladder		
2	PCR product of sample 4K	No amplification	109bp
3	PCR product of sample 5K	449bp	117bp
4	PCR product of sample 6K	420bp	117bp
5	PCR product of sample 7K	449bp	117bp
6	PCR product of sample 8K	No amplification	117bp
7	PCR product of sample 9K	449bp	No amplification
8	PCR product of sample 10K	449bp	117bp
9	PCR product of sample 11K	No amplification	No amplification
10	PCR product of sample 12K	No amplification	109bp
11	PCR product of sample 13K	No amplification	No amplification
12	PCR product of sample 14K	No amplification	No amplification
13	PCR product of sample 15K	420bp	109bp
14	Negative Control	No amplification	No amplification

Table 3

Allele and heterozygosity of TPA-25 element of Jawa-Malay population

No	Sample	Heterozygosity of TPA-25
1	43J	-ve
2	44J	+ve
3	45J	-ve/+ve
4	49J	+ve
5	50J	-ve/+ve
6	52J	-ve/+ve
7	53J	-ve/+ve
8	54J	-ve/+ve
9	55J	-ve/+ve
10	56J	-ve/+ve
11	57J	-ve
12	58J	-ve/+ve
13	59J	+ve
14	20J	-ve/+ve
15	21J	+ve
16	24J	-ve
17	25J	+ve
18	29J	-ve/+ve
19	31J	-ve/+ve
20	30J	+ve
21	32J	-ve/+ve
22	33J	+ve
23	34J	No allele
24	16J	-ve
25	17J	-ve/+ve
26	11J	No allele
27	10J	-ve/+ve
28	7J	No allele
29	6J	-ve/+ve
30	5J	-ve/+ve
31	4J	+ve
32	41J	No allele
33	42J	+ve

+ve = homozygous insertion; -ve = homozygous deletion; -ve/+ve = heterozygous allele.

Table 4

Allele and heterozygosity of TPA-25 element of Kelantan -Malay population

No	Sample	Heterozygosity of TPA-25	No	Sample	Heterozygosity of TPA-25
1	4K	-ve	28	32K	+ve
2	5K	-ve/+ve	29	35K	+ve
3	6K	-ve/+ve	30	37K	-ve
4	7K	-ve/+ve	31	39K	-ve
5	8K	No allele	32	42K	+ve
6	9K	-ve/+ve	33	44K	+ve
7	10K	+ve	34	46K	-ve
8	11K	-ve	35	50K	-ve
9	12K	No allele	36	55K	-ve/+ve
10	13K	No allele	37	81K	-ve/+ve
11	14K	-ve	38	89K	-ve/+ve
12	15K	+ve	39	91K	No allele
13	16K	-ve	40	95K	-ve
14	17K	-ve	41	97K	-ve/+ve
15	18K	-ve/+ve	42	99K	-ve/+ve
16	19K	-ve/+ve	43	101K	-ve
17	20K	-ve	44	86K	-ve
18	21K	No allele	45	57K	-ve/+ve
19	23K	-ve	46	58K	+ve
20	24K	+ve	47	63K	+ve
21	25K	-ve/+ve	48	68K	No allele
22	26K	-ve	49	70K	-ve/+ve
23	27K	-ve/+ve	50	71K	-ve
24	28K	-ve	51	72K	-ve/+ve
25	29K	-ve			
26	30K	-ve/+ve			
27	31K	No allele			

+ve = homozygous insertion; -ve = homozygous deletion; -ve/+ve = heterozygous allele.

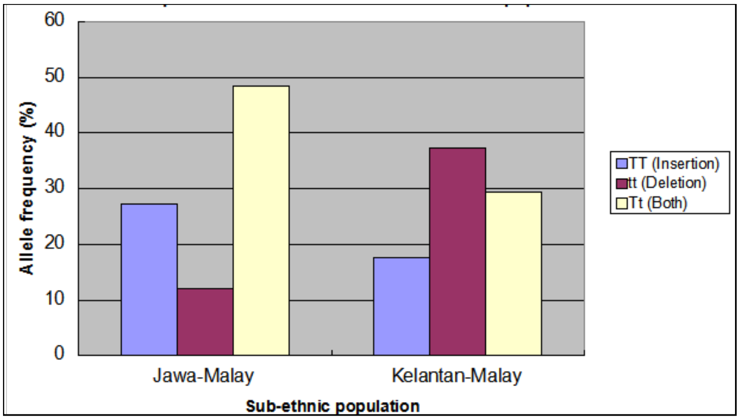


Figure 4. Percentage of allele frequency of TPA-25 (heterozygosity)

The heterozygosity and F-statistic for TPA-25 element among the two Malays sub-ethnic populations in peninsular Malaysia were calculated in F-statistic indices which had been formalized by Wright (1978) using the three different hierarchical measures of heterozygosity, H (HI, HS and HT) as shown in Table 5.

The inbreeding coefficient (FIS) assesses the global variation in individuals, relative to the variation in their subpopulations. From the results shown in Table 6, the FIS values for two sub-populations were positive which indicate that the subpopulation is deficit of heterozygotes. However, the FIS value for Kelantan-Malay population was higher than Jawa-Malay population. This data indicates that occurrence of inbreeding is higher in Kelantan-Malay population compared to Jawa-Malay population.

The overall inbreeding coefficient (FIT) assesses the variation in individuals relative to the variation in the total set of subpopulations. FIT value for the populations was 0.5359 which was within range 0 to 1, meaning the TPA-25 element had significant variation among individuals within the total population of Jawa-Malay and Kelantan-Malay studied. Fixation index, FST is the statistic used to assess the variation in the subpopulations relative to that in the total population.

Table 5

Heterozygosity value of TPA-25 element for two Malays sub-ethnics

	Alleles		Subpopulation		Total	
	p	H	P	H	P	H
Jawa-Malay	0.4848	0.7650	0.1905	0.4848		
Kelantan-Malay	0.2941	0.9135	0.1785	0.2941	0.1845	0.8393
Average Heterozygosity	$H_S =$	0.8393	$H_I =$	0.3895	$H_T =$	0.8393

Table 6

F-statistic of TPA-25 element for two Malays sub-ethnics

Subpopulation	Jawa-Malay		Kelantan-Malay
F_{IS}	0.3663		0.6781
All population			
F_{IT}		0.5359	
F_{ST}		0.0	

The value of FST for the total population was 0. As referred to qualitative guidelines for FST interpretation by Wright (1978), the FST value indicates that there is no subdivision, random mating occurring and no genetic divergence within the populations. However, the expected result for FST value was opposite from the data recorded as these two Malays sub-ethnic populations should show some genetic variation because the possibility of genetic drift happened within the population was high if referred to the FIT and FIS values. This misinterpreted data may be caused by the small scale of sample that was used in this study preventing the use of the parameter to determine the genetic differentiation of overall populations.

In the comparison with other world population, using TPA-25 element as the ancestry marker, this Alu insertion is proven as a useful tool that can reveal the patterns of human genetic diversity over the last 100,000 years (Rishishwar et al., 2015). Table 7 shows the TPA-25 frequencies in various world populations with range of frequencies within 0.3 to 0.6.

Table 7

Allele frequencies of TPA-25 in the world populations

Population	n	f	References
Europe			
Albania	60	0.557	Comas et al. (2000)
Albania Aromuns	49	0.500	Comas et al. (2000)
Andalusia	67	0.590	Comas et al. (2000)
Basque	96	0.568	Comas et al. (2000)
Catalonia	60	0.608	Comas et al. (2000)
Genova	30	0.450	Santovito et al. (2007)
Germany	70	0.514	Romualdi et al. (2002)
Valencia	101	0.556	García-Obregón et al. (2006)
North Africa			
Algeria	47	0.532	Comas et al. (2000)
NSC Tunisia	96	0.572	Bahri et al. (2008)
N. Morocco	111	0.617	Comas et al. (2000)
Sahara	58	0.397	Comas et al. (2000)
Sejnane	47	0.521	Frigi et al. (2011)
SE. Morocco	49	0.510	Comas et al. (2000)
Tunisia	48	0.604	Comas et al. (2000)
Takrouna	33	0.469	Frigi et al. (2011)
W. Morrocco	140	0.575	Comas et al. (2000)

Table 7 *(Continued)*

Population	n	f	References
South India			
Koya Dora	59	0.458	Veerraju et al. (2001)
Konda Reddi	57	0.481	Veerraju et al. (2001)
Central Patagonia			
Comodoro Rivadavia	50	0.500	Parolin et al. (2017)
Esquel	50	0.388	Parolin et al. (2017)
Puertp Madryn	50	0.620	Parolin et al. (2017)

n: number of samples; f: allele frequency

CONCLUSION

Kelantan-Malay population has genetically conserved their own special traits with low occurrence of intermarriage with other Malays sub-ethnic group and high occurrence of inbreeding within the population. Meanwhile, low inbreeding in Jawa-Malay population indicates is big genetic diversity exists for this population in peninsular Malaysia compared to Kelantan-Malay population.

Moreover, migration flows and interactions between sub-ethnic groups are suggested to be the main cause for genetic diversity occuring in both populations. This research suggests that culture and social life of these populations affect the genetic flows as nowadays, majority of them kept standard Malays lifestyle and cultures instead following their ancestry practices.

ACKNOWLEDGEMENT

This research was supported by Research and Ethics Committee of Universiti Teknologi MARA (UiTM).

REFERENCES

Ade, C., Roy-Engel, A. M., & Deininger, P. L. (2013). Alu elements: An intrinsic source of human genome instability. *Current Opinion in Virology, 3*(6), 639-645.

Asari, M., Omura, T., Oka, K., Maseda, C., Tasaki, Y., Shiono, H., … & Shimizu, K. (2012). Multiplex PCR-based Alu insertion polymorphisms genotyping for identifying individuals of Japanese ethnicity. *Genomics, 99*(4), 227-232.

Bahri, R., Esteban, E., Moral, P., & Chaabani, H. (2008). New insights into the genetic history of Tunisians: Data from Alu insertion and apolipoprotein E gene polymorphisms. *Annals of Human Biology, 35*(1), 22-33.

Bamshad, M. J., Wooding, S., Watkins, W. S., Ostler, C. T., Batzer, M. A., & Jorde, L. B. (2003). Human population genetic structure and inference of group membership. *The American Journal of Human Genetics, 72*(3), 578-589.

Batzer, M. A., & Deininger, P. L. (2002). Alu repeats and human genomic diversity. *Nature Reviews Genetics, 3*(5), 370-379.

Batzer, M. A., Kilroy, G. E., Richard, P. E., Shaikh, T. H., Desselle, T. D., Hoppens, C. L., & Deininger, P. L. (1991). Structure and variability of recently inserted Alu family members. *Nucleic Acids Research, 19*(3), 698-698.

Bellwood, P. (2007). *Prehistory of the Indo-Malaysian archipelago*. Canberra, Australia: ANU E Press.

Cherni, L., Frigi, S., Ennafaa, H., Mtiraoui, N., Mahjoub, T., & Benammar-Elgaaied, A. (2011). Human Alu insertion polymorphisms in North African populations. *Human Biology, 83*(5), 611-627.

Comas, D., Calafell, F., & Benchemsi, N. (2000). Alu insertion polymorphisms in NW Africa and the Iberian Peninsula: Evidence for a strong genetic boundary through the Gibraltar Straits. *Human Genetic, 107,* 312-319.

Cordaux, R., Hedges, D. J., Herke, S. W., & Batzer, M. A. (2006). Estimating the retrotransposition rate of human Alu elements. *Gene, 373,* 134-137.

Cordaux, R., Srikanta, D., Lee, J., Stoneking, M., & Batzer, M. A. (2007). In search of polymorphic Alu insertions with restricted geographic distributions. *Genomics, 90*(1), 154-158.

Ennafaa, H., Amor, M. B., Yacoubi-Loueslati, B., Khodjet El Khil, H., Gonzalez-Perez, E., Moral, P., & Elgaaied, A. (2006). Alu polymorphisms in Jerba Island population (Tunisia): comparative study in Arab and Berber groups. *Annals of Human Biology, 33*(5-6), 634-640.

Frigi, S., Ennafaa, H., Ben Amor, M., Cherni, L., & Ben Ammar-Elgaaied, A. (2011). Assessing human genetic diversity in Tunisian Berber populations by Alu insertion polymorphisms. *Annals of Human Biology, 38*(1), 53-58.

García-Obregón, S., Alfonso-Sánchez, M. A., Pérez-Miranda, A. M., Vidales, C., Arroyo, D., & Pena, J. A. (2006). Genetic position of Valencia (Spain) in the Mediterranean basin according to Alu insertions. *American Journal of Human Biology: The Official Journal of the Human Biology Association, 18*(2), 187-195.

Hugo, G. (1993). Indonesian labour migration to Malaysia: Trends and policy implications. *Southeast Asian Journal of Social Science, 21*(1), 36-70.

Jacq-Hergoualc'h, M. (2018). *The Malay Peninsula: Crossroads of the maritime silk road (100 BC-1300 AD).* Leiden, The Netherlands: Brill.

Keita, S. O. Y. (2010). Biocultural Emergence of the Amazigh (Berbers) in Africa: Comment on Frigi et al (2010). *Human Biology, 82*(4), 385-393.

Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., … & FitzHugh, W. (2001). Initial sequencing and analysis of the human genome. *Nature, 409*(6822), 860-921.

Moncer, W. E., Bahri, R., Esteban, E., Abdenni-Guenounou, B., Moral, P., Chibani, J. B., & Chaabani, H. (2011). Research of the origin of a particular Tunisian group using a physical marker and Alu insertion polymorphisms. *Genetics and Molecular Biology, 34*(3), 371-376.

Omar, A. H. (Ed.). (2004). *The encyclopedia of Malaysia: languages and literature (Vol. 9).* Singapore: Archipelago Press.

Parolin, M. L., Zanetti, D., Calò, C. M., Esteban, E., Avena, S. A., Carnese, F. R., & Moral, P. (2017). Migrations, admixture and genetic diversity in Central Argentinian Patagonia: analysis of autosomal Alu polymorphisms. *Journal of Basic and Applied Genetics, 28*(2), 29-42.

Ray, D. A., Walker, J. A., Hall, A., Llewellyn, B., Ballantyne, J., Christian, A. T., … & Batzer, M. A. (2005). Inference of human geographic origins using Alu insertion polymorphisms. *Forensic Science International, 153*(2), 117-124.

Rishishwar, L., Villa, C. E. T., & Jordan, I. K. (2015). Transposable element polymorphisms recapitulate human evolution. *Mobile DNA, 6*(1), 1-12.

Roewer, L., Kayser, M., de Knijff, P., Anslinger, K., Betz, A., Caglia, A., & Kärgel, H. J. (2000). A new method for the evaluation of matches in non-recombining genomes: application to Y-chromosomal short tandem repeat (STR) haplotypes in European males. *Forensic Science International, 114*(1), 31-43.

Romualdi, C., Balding, D., Nasidze, I. S., Risch, G., Robichaux, M., Sherry, S. T., ... & Barbujani, G. (2002). Patterns of human diversity, within and among continents, inferred from biallelic DNA polymorphisms. *Genome Research, 12*(4), 602-612.

Santovito, A., Selvaggi, A., Cervella, P., Castellano, S., Bigatti, M. P., Sella, G., & Delpero, M. (2007). Polymorphic Alu insertions in five North-West Italian populations. *American Journal of Human Biology: The Official Journal of the Human Biology Association, 19*(4), 589-592.

Shamsul, A. B. (2001). A history of an identity, an identity of a history: the idea and practice of 'Malayness' in Malaysia reconsidered. *Journal of Southeast Asian Studies, 32*(03), 355-366.

Slatkin, M. (1995). A measure of population subdivision based on microsatellite allele frequencies. *Genetics, 139*(1), 457-462.

Tishkoff, S. A., & Kidd, K. K. (2004). Implications of biogeography of human populations for 'race' and medicine. *Nature Genetics, 36,* S21-S27.

Veerraju, P., Rao, T. V., Lakshmi, N., Reshmi, S., Dey, B., & Majumder, P. P. (2001). Insertion/deletion DNA polymorphisms in two south Indian tribal populations. *International Journal of Human Genetics, 1*(2), 129-132.

Wheatley, P. (1961). *The Golden Khersonese: Studies in the historical geography of the Malay Peninsula before AD 1500*. Kuala Lumpur, Malaysia: University of Malaya Press.

Wright, S. (1978). V*ariability within and among natural populations Vol. 4: Evolution and the genetics of populations: a treatise in four volumes.* Chicago, IL: University of Chicago Press.

Pertanika J. Sci. & Technol. 27 (4): 1589 - 1602 (2019)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Effect of Diet Regime on the Development and Survival of *Aedes albopictus* (Skuse) (Diptera: Culicidae)

Wan Nurul Fatin Wn Maamor[1], Nazri Che Dom[1,2]*, Megat Azman Megat Mokhtar[1] and Siti Nazrina Camalxaman[1,2]

[1]*Faculty of Health Sciences, Universiti Teknologi MARA, 42300 UiTM, Puncak Alam Selangor, Malaysia*
[2]*Integrated Mosquito Research Group (I-MeRGe), Universiti Teknologi MARA, 42300 UiTM, Puncak Alam, Selangor*

ABSTRACT

A study was conducted to identify the responses of *Aedes albopictus* to different diet regime towards the development of juvenile and adult mosquitoes. Fish pellet was selected as standard diet in order to study the effect of diet regime on the development of *Ae. Albopictus*. Four different diet regimes (1.0 mg, 0.6 mg, 1.0 mg an1.6 mg) were tested on 50 eggs of *Ae. albopictus* under laboratory conditions. Juvenile development until adult emergence was observed and recorded. Results indicated that the time taken to mature the mosquitoes was significantly affected by the diet regime. Furthermore, juvenile body size and adult wing size of *Ae. albopictus* were found to be greatly affected by diet regime exposed during juvenile stages. In summary, an increase of diet regime resulted in the decrease of developmental time and an increase in juvenile body size and adult wing size.

Keywords: Aedes, Ae. albopictus, diet, juvenile size, wing size.

ARTICLE INFO

Article history:
Received: 1 September 2018
Accepted: 4 June 2019
Published: 21 October 2019

E-mail addresses:
wnurulfatinwmaamor@gmail.com (Wan Nurul Fatin Wn Maamor)
nazricd@salam.uitm.edu.my (Nazri Che Dom)
meguitm@gmail.com (Megat Azman Megat Mokhtar)
sitinazrina@salam.uitm.edu.my (Siti Nazrina Camalxaman)
*Corresponding author

INTRODUCTION

Dengue fever (DF) is endemic and growing a public health problem which affecting the human health. (Getachew et al., 2015; Sairi et al., 2016; Rozilawati et al., 2017). Development activities, population growth, and increased individual movements have contributed to the emergence and reemergence of the disease (Chen et al.,

ISSN: 0128-7680
e-ISSN: 2231-8526

2009; Getachaw et al., 2015). Dengue viruses are the causative agents of DF which poses a serious health issues (Getachew et al., 2015; Vijayakumar et al., 2014). The viruses are transmitted through the bites of female *Aedes* mosquitoes. The main vectors in the transmission of dengue fever infection are *Aedes albopictus* (Skuse), otherwise known as known as the "Asian tiger mosquito" and *Aedes aegypti* (Linnaeus). *Ae. albopictus* prefers to breed in natural habitat and condition with more vegetation such as trees holes, rock pools and other natural sites (Rao et al., 2011; Ishak et al., 2015). However, deforestation activities, climate change and the increase in global trade have enabled this species to adapt with different environment. Few studies have observed the presence of this species in domestic and semi-domestic artificial container habitats (Dom et al., 2016a; Rao et. al., 2011).

Many studies have documented that the rate of juvenile development and its survival are influenced by biotic and abiotic factor of the environment (Couret et al., 2014; Dom et al., 2016b; Koenraadt & Harrington, 2008; Legros et al., 2009; Manorenjitha & Zairi, 2012; Rowbottom et al., 2015; Teng & Apperson, 2000). Limitation of food supply during juvenile cycle is a probable factor that may affect the adult survivorship and can contribute the body size of the adult mosquitoes (Jannat & Roitberg 2013). A recent study conducted by Araujo et al. (2012) stated that the survival of larvae increased with increasing food supply. This study also highlighted that improper food supply might delay the development of the larvae and thus decreased their survival rate during the fourth instar stages. Therefore, it is important to assess the influences of food supply on juvenile development in order to understand the dynamic of *Aedes* mosquito ecology; justifying the rationale of this study.

There is a group of researcher conducting a study on diet regime towards the development of *Aedes* mosquitoes. There are six diet ingredients used to study on the development of the *Aedes* mosquito species which are (a) beef liver powder, tuna meal, vitamin mix in water, (b) liver powder, (c) dog biscuit, beef liver, milk powder, (d) tuna meal, low heat desiccated non-defatted solid, Argentines beef liver powder, vitamin, (e) fish pellet and (f) dog food, dried beef liver, yeast, and milk powder. However, most of study used fish pellet formulated with necessary amino acid, protein, fiber, minerals and vitamin as their main ingredients and nutrition supply for development of juvenile mosquito (Arrivillaga & Barrera, 2004; Araujo et al., 2012; Couret et al., 2014; Manorenjitha & Zairi, 2012; Jannat & Roitberg, 2013; Jong et al., 2016; Yoshioka et al., 2012)

Study conducted by Araujo et al. (2012) had used fish pellet from TetraMin tropical flaskes-spectrum brands as their main food ingredients. This food contained nutritionally balanced with selected vitamins, mineral and trace elements that required for juvenile development. Besides that, this study supplied the food regime (low, medium, high) based on the food amounts consumed by other species of culicidae. They concluded that the development time decreased with the increasing of food regime which the first and fourth instar took longer time to develop when immature supplied with low and medium food amounts. Besides that, a study conducted in laboratory condition also found that development time of *Aedes albopictus* juvenile decreased with increasing of the diet level. They concluded that the longest pupation time was when larvae treated with 2.0 mg (14 days) diet regime while shortest time of pupation when larvae treated with 20.0 mg (8 days) (Yoshioka et al., 2012).

In addition, Arrivillaga and Barrera (2004) carried out a study that focussed on the food limitation in water storage container on development of *Ae. aegypti*. In this study, they used liver powder as their diet ingredients that supplied the main nutrition (carbohydrates, vitamin, mineral) that required for immature development. They found that immature supplied with lowest food level (0.01 mg / larval / day) showed adult emergence on day nineteen while immature reared under highest food level (1.6 mg / larval / day) emerged on day nine and produce adult with mass 2.338 mg. In 2016, study in Malaysia by Jong and colleagues focus on the minimum feeding regime required to produce viable and competitive adult *Ae. albopictus*. They had used combination of dog food, dried beef liver, yeast, and milk powder as food ingredient and stated that the duration of each immature stage was not affected by different feeding regimes except at the fourth instar. However, duration of the first instar was slightly prolonged with the reduction in feeding regime. Besides that, this study had followed standard feeding regime provided by VCRU in Universiti Sains Malaysia for rearing *Aedes* larvae which was 0.75 mg / larva / day as the control treatment.

Apart from that, three study conducted had focussed on the food and density influence on the juvenile development (Couret et al., 2014; Jannat & Roitberg, 2013; Manorenjitha & Zairi, 2012). This study observed effect of the overcrowding among the larvae to get the nutrition sources for development. As the finding, larvae reared under (1 mg) sub-optimal larval food condition had long development time (9 days) and for larvae supplies with optimal food (100 mg) took seven days to pupation (Manorenjitha & Zairi, 2012). Besides that, juvenile treated with high food treatment had low mortality rate which was twenty percent. While, juvenile with low food sources had highest mortality rate which is hundred percent (Jannat & Roitberg, 2013). Couret et al. (2014) had used combination of beef liver powder, tuna meal, vitamin and water as the main ingredient for food. This food contained all the important nutrition for the juvenile development including protein, amino acid, carbohydrates, vitamin and mineral and the diet regime was added daily based on the number of larvae alive in order to ensure amount of food offered per larva per day was kept constant throughout the juvenile development. However, high density of larvae with low food resources caused lowers development rates. There are major gap on the study of influence of diet on development of *Aedes* mosquito. The data on the effect of the diet on the life development of *Aedes* mosquito is still limited in this country. Thus, it is very important to study the effect of diet on the development of this vector mosquito in the local environment using local strain in order to understand on the biology of mosquito.

MATERIALS AND METHODS

This study utilized an experimental study design to investigate the effects of different diet regime on the development of *Ae. albopictus*. The comparative study was conducted by introducing laboratory strains of *Ae. albopictus* into four different diet regimes under laboratory setting. The experiment was completely randomized in fractional 4 x 2 replication (8 treatments). Independent variables consisted of four diet regimes (0.1 mg, 0.6 mg, 1.0 mg, and 1.6 mg) and two replicates. The data collected for each experiment were tabulated and analyzed using descriptive and statistical values by ANOVA. A laboratory strain F185 generation was used from Vector Control and Research Unit of Universiti Sains Malaysia to get F187 generation. Colonies were established and maintained in a thermostatically controlled insectary at 25 ± 2^{o}C and 70 ± 10% relative humidity (RH) and 12:12 hours light: dark cycle. The rearing process was observed on a daily basis until F187 generation.

Pilot study was conducted in order to identify standard diet ingredient on juvenile development. This pilot study focused on observation of juvenile development until pupation. Three types of ingredient were used in this pilot study which are; (i) liver powder (D_1), (ii) fish pellet (D_2) and (iii) combination of cat biscuit, milk powder, yeast, chicken liver (D_3). A total of 50 first larvae was placed into the plastic container with 300 ml of deionized water. Each day 0.75 mg of (D_1) liver powder was added based on survival larvae. Cohorts were treated with a total larval diet of 22.5 mg (0.75 mg x 50 larvae) on day one. The larval diet was added daily until all larvae pupated. The procedures were replicated three times and repeated for other diet ingredient. The optimized diet ingredient that had high development rate and low morality rate from this pilot study was used for experimental phase. In this study fish pellet was selected as standard diet in order to study the effect of diet regime on the development of *Ae. albopictus*.

The observation on development on the *Ae. albopictus* with different diet regime (0.1, 0.6, 1.0, 1.6 mg / larvae / day) was conducted. Larval diet was offered once to each cohort throughout the experiment instead adding it from day one to day five, when the larvae were expected to pupate. Optimized diet ingredients from the pilot study were weighed using analytical electronic balance based on the calculation (mg x 50 larvae x 5 day), which were F1 (25 mg), F2 (150 mg), F3 (250 mg) and F4 (400 mg). For example, cohort F2 was treated with the feeding regime 0.6 mg / larvae / day were given a total diet of 150 mg (0.6 mg x 50 larvae x 5 days). Four plastic containers (10 x 10 x 10cm) were label with F1, F2, F3 and F4. A total of 50 eggs were submerged in plastic container with different label and the development was observed. Different diet regime was introduced to the egg in the water container. The experiments were conducted in replicates, twice for every diet regime. The growth and survival of juvenile until adult emergence were monitored and recorded daily. The procedures were repeated for different diet regime is summarized in Figure 1.

Data collection was initiated when the first instar was introduced to a different diet. Visual observation was conducted every day at the same time (4.00 pm) to monitor the development of juvenile until adult emergence. Stages of the juvenile, time of pupation and days of adult emergence were recorded. The survival rate of each treatment was quantified by recording all survival juvenile and adult on a daily basis per container and all dead larvae were removed whenever they were observed. Data on 50% pupation was expressed by the day that half of the instar in the water container pupate. Juvenile body size was observed using standardized digital of random individuals every 24 hour starting on first instar. Measuring the size of juvenile was using Dino lite with open-access software DinoXcope version 1.18 (Madzlan et al., 2018). Adult emergence was collected and placed in a labeled universal bottle. This bottle was placed in a refrigerator at 4^{o}C for two hours after adult collection in order to kill the adult. Wing was removed and mounted on microscope slide. These were then digitally scanned and measured under Dino-Lite using an open-access software DinoXcope version 1.18.

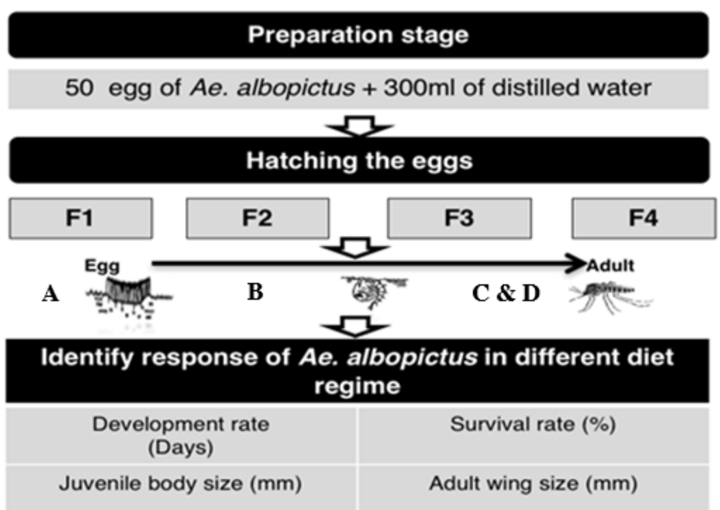


Figure 1. Flow process of the study on different diet regime.

*Note: F1 (0.1 mg), F2 (0.6 mg), F3 (1.0 mg), F4 (1.6 mg). Several aspect on the development trend were monitored which are; (A) eggs hatching in days and percentage, (B) Fifty percent pupation in day, (C) fifty percent adult emergence and (D) hundred percent adult emergence

The wing size measurement of adult mosquito was also conducted. The dead mosquitoes were kept in the freezer to allow a measurement of wing lengths (Blackmore & Lord, 2000). The wing length was defined as the linear distance from the axillary incision to the apical margin excluding the fringe scales (Loetti et al., 2011; Neira et al., 2014; Schneider et al., 2004). To measure the length, the wings were detached proximal to the axillary notch, and mounted on a microscope slide and covered with a cover slip (Schneider et al., 2004; Vidal & Suesdek, 2011). Only one wing of each adult was removed which was usually the left wing, unless it was damaged (Reiskind & Zarrabi, 2012).

In this study, the diameters of wing shapes measurement were used and this technique is unconventional to represent wing size/wing length (Figure 2). In mosquito studies (Culicidae), the wing is widely used for morphometric comparison because of its two-dimensional shape and because it contains veins that encompasses natural anatomical landmarks that are ideal for measurement (Lorenz et al., 2017). By measuring the shape, it is possible to observe the variation of wing shape. Besides that, this measurement technique is fast, inexpensive and simple which makes possible to compare shape with minimum interference from different size.

Figure 2. Measurement of the diameter of wing shape outline by red color of *Ae. albopictus*

The response of *Ae. albopictus* on different diet regime was determined by observing the response of eggs from hatching until adult emergence on four different diet regimes. Several aspect on the development trend were monitored which are; (A) eggs hatching in days and percentage, (B) Fifty percent pupation in day, (C) fifty percent adult emergence and (D) hundred percent adult emergence. For this study, the data on the effects of diet regime on the development of juvenile *Aedes albopictus* were analyzed with analysis of variance (ANOVA) through Statistical Package of Social Science (SPSS) software.

Effect of Diet on the Development Time and Survival of *Ae. albopictus*

Generally, the development time decreased with increasing of diet regime (Table 1). The day of the first egg hatching in different diet regime was recorded and labeled as stage A.

Eggs exposures to the highest diet regime showed the shortest period for hatching (1.6 mg: 2.00 ± 0.00 days). In contrary, exposure to low diet regime (1.0 mg, 0.6 mg and 0.1 mg) showed longer period of egg hatching (13.00 ± 0.00 days; 3.17 ± 0.75; 3.83 ± 0.41 days respectively). Data on stages B was collected by recording the day that 50% of the larvae changed to pupae. It was observed that larvae supply with lowest diet regime took longest time for pupation (0.1 mg: 9.17 ± 0.75 days). While, other diet regimes showed a shorter time for pupation (1.6 mg: 7.50 ± 0.55 days; 1.0 mg: 8.00 ± 0.63; 0.6 mg: 8.50 ± 0.55 days).

Adult emergence was observed until 50% of the pupae emergence into adult and the data was collected to represent stages C. Fifty percent of adult emergence range from 8.83 to 11.00 days. Groups supplied with highest diet regime (1.6 mg) were found to be the fastest to emerge (8.83 ± 0.41 days) as compared to other diet regimes (1.0 mg: 9.67 ± 0.52 days; 0.6 mg: 10.33 ± 0.52; 0.1 mg: 11.00 ± 0.63 days). Stages D was recorded when the entire juvenile emerged as adult mosquitoes in different diet regimes. During this stage, the group with lowest diet regime (0.1 mg) took the longest time to emerge (14.83 ± 0.41 days) compared to other group of diet regime. As for the increase of diet regime, the duration for full adult emergence also decreased (0.6 mg: 14.00 ± 0.63 days; 1.0 mg: 12.83 ± 0.41; 1.6 mg: 11.50 ± 0.84 days). The ANOVA results was statistically significant, indicating that the development duration was influenced by diet regime, $F(3,20)$, $p= < 0.05$. Post hoc analysis with Tukey's HSD (using an α of 0.05) revealed that juvenile supply with lowest diet regime (M= 14.83, SD=0.41) had significantly longer the duration of development followed by diet regime 0.6 mg (M= 14.00, SD=0.63), diet regime 1.0 mg (M= 12.83, SD=0.41) and diet regime 1.6 mg (M= 11.50, SD=0.84).

Table 1

Duration of development of Ae. albopictus feed with four different diet regimes (0.1 mg, 0.6 mg, 1.0 mg and 1.6 mg) under controlled conditions (25 ± 2oC and 70 ± 10% relative humidity (RH) and 12:12 hours light: dark cycle).

Stages	0.1 mg	0.6 mg	1.0 mg	1.6 mg	*p-values
A	3.8±0.41	3.2±0.75	3.0±0.00	2.0±0.00	<0.05
B	9.2±0.75	8.5±0.55	8.0±0.63	7.5±0.55	<0.05
C	11.0±0.63	10.3±0.52	9.7±0.52	8.8±0.41	<0.05
D	14.8±0.41	14.0±0.63	12.8±0.41	11.5±0.84	<0.05

Note:Stages A: Eggs hatching; B: Fifty percent of larvae change to pupae; C: Fifty percent of pupae change to adult and D: Pupae change to adult.
*Data are represented as Means ± SD. *There has significant difference found between duration of development of Ae. albopictus exposure with different diet regime p=< 0.05 based on one-way ANOVA.*

In addition, the survival rate of *Ae. albopictus* was also affected by different diet regime ($p < 0.05$) (Table 2), which observed a decrease in numbers with increasing of diet regime. Group that was fed with highest diet regime (1.6 mg) showed the lowest survival rate in all stages (stage A: 79.00 ± 1.41%; stage B: 62.50 ± 3.53 %; stage C: 50.00 ± 0.00%; stage D: 37.00 ± 1.41%). However, the group exposed to lowest diet regime showed the highest survival rate during juvenile and adult stages (stage A: 96.50 ± 0.71%; stage B: 87.50 ± 0.71%; stage C: 84.50 ± 6.36%; stage D: 71.50 ± 0.71%). The ANOVA was statistically significant, indicating that survival rate was influenced by diet regime, $F(3,20)$, p= < 0.05. Post hoc analysis with Tukey's HSD (using an α of 0.05) revealed that juvenile supply with lowest diet regime (M= 71.50, SD=0.71) had significantly highest survival rate followed by diet regime 0.6 mg (M= 68.5, SD=0.71), diet regime 1.0 mg (M= 65.5, SD=0.71) and diet regime 1.6 mg (M= 37.0, SD=1.41). Generally, in duration of development aspect, the development time decreased with increase of diet regime. However, the survival rate decreased with an increase of diet regime.

Table 2

The survival rate of Ae. albopictus feed with four different diet regimes (0.1 mg, 0.6 mg, 1.0 mg and 1.6 mg) under controlled conditions (25 ± 2°C and 70 ± 10% relative humidity (RH) and 12:12 hours light: dark cycle)

	Survival rate (%) ± SD, (*n*= 50/trials)				
	Food amount				
Stages	0.1 mg	0.6 mg	1.0 mg	1.6 mg	p-values
A	96.5±0.71	92.5±0.71	87.0±1.41	79.0±1.41	<0.05
B	87.5±0.71	82.5±0.71	77.5±0.71	62.5±3.53	<0.05
C	84.5±6.36	72.5±0.71	69.5±0.71	50.00±0.00	<0.05
D	71.5±0.71	68.5±0.71	65.5±0.71	37.0±1.41	<0.05

Note: Stages A: Eggs hatching; B: Fifty percent of larvae change to pupae; C: Fifty percent of pupae change to adult and D: Pupae change to adult.

*Data are represented as Means ± SD. *There has significant difference found between duration of development of Ae. albopictus exposure with different diet regime p=< 0.05 based on one-way ANOVA.*

Previous studies have described that high diet condition correlates with shorter *Ae. albopictus* development time (Araujo et al., 2012; Manorenjitha & Zairi, 2012; Yoshioka et al., 2012). In the present study, all the groups exposed to different diet regime were able in completing the development cycles until adult. However, the shortest development period of *Ae. albopictus* (from eggs to pupae) was observed at highest diet regime. On the other hand, the longest development duration (14.83 days) was observed at lowest diet regime, probably as a result of insufficient energy to complete development (Arrivillaga & Barrera, 2004). Poor diet causes an extended development time and since immature spend 25% of their biomass and average moulting, mosquito larvae must require enough food supply for ecdysis (Araujo et al., 2012). In addition, this experiment showed that, survival rate decreased with an increase of diet regime. This result in line with previous study by Jannat and Roitberg (2013), where the authors reported that survival rate increased with decrease of food sources. Observation during the experiment showed that water with high diet regime produced scum and smell. The scum were produced from the gradual decomposition of organic matter (diet) and deposited as sediment and smell were produced due to the stagnant water condition which all the microorganisms using up all of the available oxygen.

Effect of Diet on the Juvenile Body Size of *Ae. albopictus*

The mean of juvenile body size of *Ae. albopictus* was significantly changed with diet regime (Figure 3). Adult wing size increased with the amount of food supplied during juvenile stages. Mean of juvenile body size varied between 10.94±0.35 mm to 13.07±0.55 mm. Juvenile body size of *Ae. albopictus* raised under a diet regime 0.1 mg/larva/day and below was shorter than 11.00 mm, while those under diet regime more than 0.1 mg (0.6, 1.0, 1.6 mg/larva/day) were longer than 11.00 mm. The ANOVA was statistically significant, indicating that juvenile body size was influenced by different diet regime, F (3, 8), p= 0.003. Post hoc analysis with Tukey's HSD (using an α of 0.05) revealed that juvenile supply with lowest diet regime (M= 10.94, SD=0.35) had significantly lower the adult wing size followed by diet regime 0.6 mg/larva/day (M= 11.40, SD=0.54), diet regime 1.0 mg/larva/day (M= 12.43, SD=0.54) and diet regime 1.6 mg/larva/day (M= 13.07, SD=0.55 juvenile body size). As predicted, juvenile body size increased proportionally with increase in diet regime with the highest diet regime (1.6 mg) resulting in the largest mean juvenile body size. While, lowest diet regime (0.6 mg) produced smaller juvenile body size.

The differences in adult wing sized of *Ae. albopictus* were noted between group supply with different diet regime from 0.1 mg to 1.6 mg. The mean adult wing size between four different diet regimes (0.1 mg, 0.6 mg, 1.0 mg, 1.6 mg) was found significantly different (ANOVA, F= (3,20), p=0.004) **(**Figure 4A**)**. Adult wing size from the higher fed group (1.6 mg) was found to be large as compared to adult wing size from lower fed group (0.1 mg) which was smaller in size **(**Figure 4B**)**. Wing size has been commonly used as a morphometric measurement of body size and weight of mosquito. Result indicated that different diet regimes correlated with adult wing size. Each group of diet regime produced adult wing size more than 11.00 mm. This result is in agreement with a previous study by Jong et al. (2016), which stated the adult wing size larger with an increase of diet regime. Besides that, additional data juvenile body size was observed and measured in this experiment to identify their significance with different diet regimes.

Wing size has been commonly used as a morphometric measurement of body size and weight of mosquito. Result obtained in this study indicated that different diet regime correlated with adult wing size. Each group of diet regime produce adult wing size more than 11.00 mm. This result is similar to a previous study conducted by Jong et al. (2016) which stated the adult wing size was larger as increase of diet regime. Additional data juvenile body size was observed and measure in this experiment to identify their significant with different diet regime. Furthermore, juvenile body size increased proportionally with increase in diet regime with the highest diet regime (1.6 mg) resulting in the largest mean juvenile body size. While, lowest diet regime (0.6 mg) produced smaller juvenile body size.

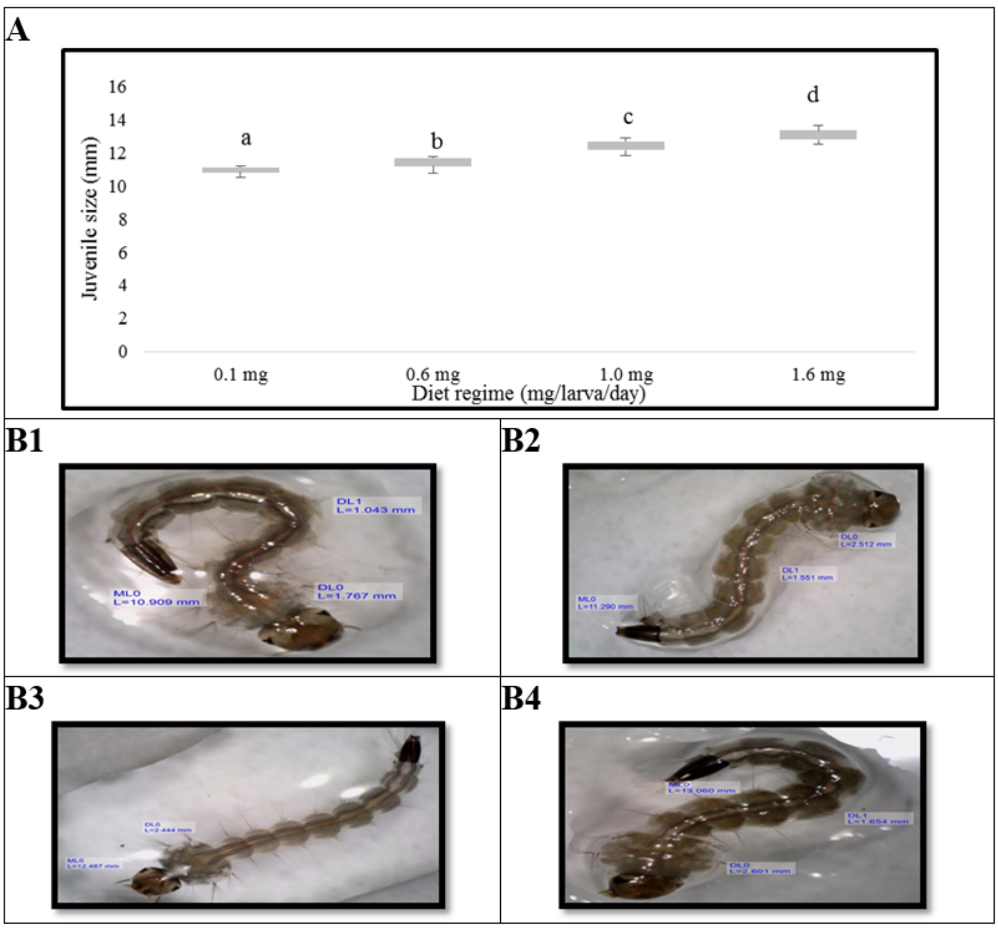


Figure 3. Effect of diet on juvenile body size of *Ae. albopictus.* (A) Data show juvenile size (mm) of *Ae. albopictus* under different diet regime (mg/larva/day). *Different letter (a, b, c, d) represent significant different between column. (B) Size of juvenile supply with different diet measure under Dino-lite.

* Note Juvenile exposure with different diet are labeled with B1,B2,B3 and B4 for 0.1 mg, 0.6 mg, 1.0 mg and 1.6 mg.

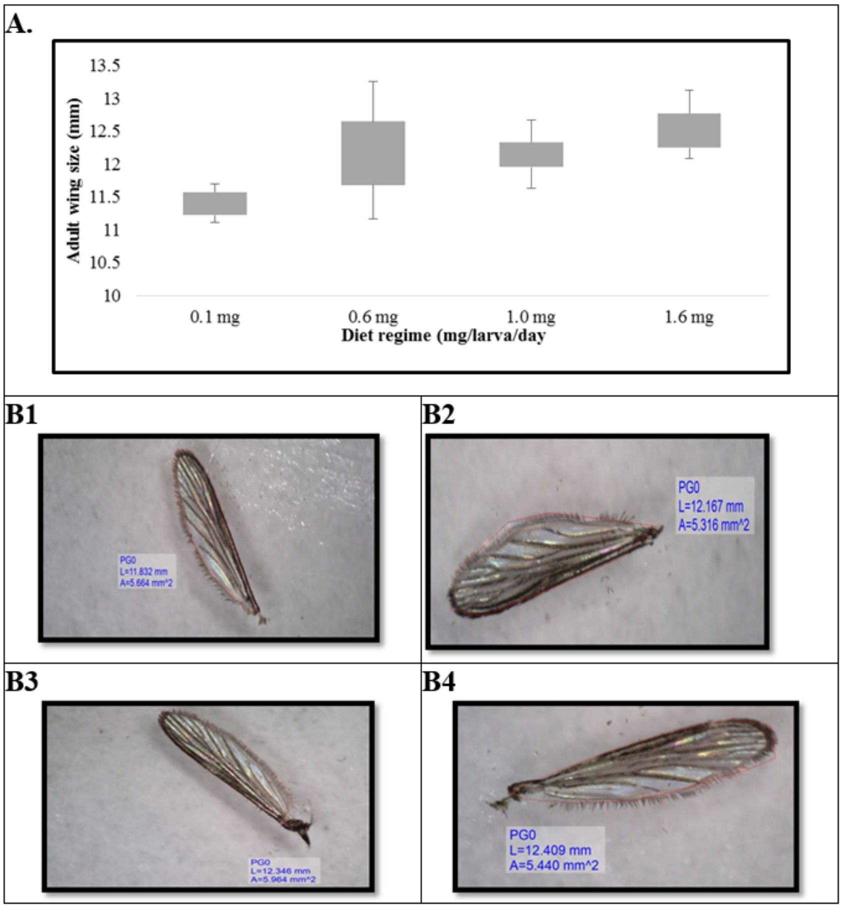


Figure 4. Effect of diet on the morphology of Ae. albopictus. (A) Box plot show a size of adult wing (mm) of Ae. albopictus under different diet regime (mg/larva/day). Shows adult wing size that supply with different diet during juvenile and measure under Dino-lite * Note Juvenile exposure with different diet are labeled with B1,B2,B3 and B4 for 0.1 mg, 0.6 mg, 1.0 mg and 1.6 mg.

Nutrition is one of the factors that influences the development and survival of mosquitoes (Arrivillaga & Barrera, 2004; Couret et al., 2014; Jong et al., 2016). Supply of the nutrition during juvenile stages affects the development rate from egg until adult stage (Jong et al., 2016 ; Manorenjitha & Zairi, 2012). The result showed wide phenotypic variation in development duration, survival rate, hatching rate, juvenile size and adult wing size to *Ae. albopictus* that reared with fish pellet under different diet regime in laboratory. The range of diet regime used in this experiment covered a wide range of possible feeding conditions in natural environment.

The rearing of *Aedes* mosquitoes is complex and demanding for several reasons. *Aedes* larvae are affected by temperature, density and available nutrition (Dom et al., 2016a; Madzlan et al., 2016; Madzlan et al., 2018). Our results showed that the fish pallet ingredient impacted the hatching rates, development times and survival of *Aedes albopictus*. Therefore this finding proposes delicate screening work and precise methods to be used for controlled laboratory strain. The advantage of this ingredient is it is readily available, convenient to treat, easy to preserve and cheap.

This study was designed purposely to determine the impact of diet regime on the development of local *Ae. albopictus*. The limitation of this study is that it only focused on few aspect on the development trend such as hatching period, pupation period and adult emergence without differentiating the developmental stages of the larval according to the stages of mosquito instar. Therefore, it is quite difficult to conclude that the effects of diet regime are related to the development period of *Aedes albopictus* due to non-consideration of stage of mosquito instar.

CONCLUSION

An understanding of the natural factors that regulate natural populations of *Ae. albopictus* mosquitoes can improve control and reduce the incidence of dengue fever cases. In general, diet regime and available breeding water container will encourage the development of *Ae. albopictus* and also reduce their development time for adult emergence. Thus, studies focusing on the monitoring effect of diet regime on the development of mosquito are essential to controlling the transmission of the dengue disease. The findings show that a higher amount of diet regime can shorten the development rate of the mosquitoes. The outcome of this study can be utilized as baseline data to provide further information on the population dynamic of the mosquito.

ACKNOWLEDGEMENTS

The authors sincerely thank to the contributer of research funding from Universiti Teknologi MARA (UiTM) - (600-IRMI/DANA 5/3/LESTARI (0093/2016), and Ministry of Higher Education (MOHE) Malaysia.

REFERENCES

Araujo, M. D. S., Gill, L. H. S., & e-Silva, A. D. (2012). Larval food quantity affects development time, survival and adult biological traits that influence the vectorial capacity of Anopheles darlingi under laboratory conditions. *Malaria Journal, 11*(261), 1-9.

Arrivillaga, J., & Barrera, R. (2004). Food as a limiting factor for *Aedes aegypti* in water storage containers. *Journal of Vector Ecology, 29*(1), 11-20.

Blackmore, M. S., & Lord, C. C. (2000). The relationship between size and fecundity in *Aedes albopictus. Journal of the Society for Vector Ecology, 25*, 212-217.

Chen, C. D., Lee, H. L., Stella-Wong, S. P., Lau, K. W., & Sofian-Aziru, M. (2009). Container survey of mosquito breeding sites in a university campus in Kuala Lumpur, Malaysia. *Dengue Bulletin, 33,* 187-193.

Couret, J., Dotson, E., & Benedict, M.Q. (2014). Temperature, larval diet, and density effects on development rate and survival of *Aedes aegypti* (Diptera: Culicidae). *PLoS ONE, 9*(2), 1-9.

Dom, N. C., Madzlan, M. F., Hasnan, S. N. A., & Misran, N. (2016a). Water quality characteristics of dengue vectors breeding containers. *International Journal of Mosquito Research, 3*(1), 25-29.

Dom, N. C., Madzlan, M. F., Yusoff, S. N. N., Ahmad, A. H., Ismail, R., & Camalxaman, S. N. (2016b). Profile distribution of juvenile Aedes species in an urban area of Malaysia. *Transactions of the Royal Society of Tropical Medicine and Hygiene, 110*(4), 237-245.

Getachew, D., Tekie, H., Gebre-Michael, T., Balkew, M., & Mesfin, A. (2015). Breeding sites of *Aedes aegypti*: potential dengue vectors in Dire Dawa, East Ethiopia. *Interdisciplinary Perspectives on Infectious Diseases, 2015,* 1-8.

Ishak, I. H., Jaal, Z., Ranson, H., & Wondji, C. S. (2015). Contrasting patterns of insecticide resistance and knockdown resistance (kdr) in the dengue vectors *Aedes aegypti* and *Aedes albopictus* from Malaysia. *Parasites and Vectors, 181*(8), 1-13.

Jannat, K. N. E., & Roitberg, B. D. (2013). Effects of larval density and feeding rates on larval life history traits in *Anopheles gambiae* spp (Diptera: Culicidae). *Journal of Vector Ecology, 38*(1), 120-125.

Jong, Z. W., Kassim, N. F. A., Naziri, M. A., & Webb, C. E. (2016). The effect of inbreeding and larval feeding regime on immature development of *Aedes albopictus*. *Journal of Vector Ecology, 42*(1), 105-112.

Koenraadt, C. J. M., & Harrington, L. C. (2008). Food as a limiting factor for *Aedes aegypti* in water-storage containers. *Journal of Medical Entomology, 45*(1), 28-35.

Legros, M., Llyord, A., Huang, Y., & Gould, F. (2009). Density dependent intraspecific competition in the larval stage of *Aedes aegypti* (Diptera: Culicidae): revisiting the current paradigm. *Journal of Medical Entomology, 46*(3), 409-419.

Loetti, V., Schweigmann, N., & Burroni, N. (2011). Development rates, larval survivorship and wing length of *Culex pipiens* (Diptera: Culicidae) at constant temperatures. *Journal of Natural History, 45*(35-36), 2203-2213.

Lorenz, C., Almeida, F., Almeida-Lopes, F., Louise, C., Pereira, S. N., Petersen, V., & Suesdek, L. (2017). Geometric morphometrics in mosquitoes: What has been measured? *Infection, Genetics and Evolution, 54,* 205-215.

Madzlan, M. F., Dom, N. C., Zakaria, N., Hasnan, S. N. A., Tiong, C. S., & Camalxaman, S. N. (2018). Profiling of dengue vectors breeding habitat at urban residential areas in Shah Alam, Malaysia. *Serangga, 22*(2), 185-216.

Madzlan, F., Dom, N. C., Tiong, C. S., & Zakaria, N. (2016). Breeding characteristics of aedes mosquitoes in dengue risk area. *Procedia-Social and Behavioral Sciences, 234,* 164-172.

Manorenjitha, M. S., & Zairi, J. (2012). Nutrition and overcrowding effects on larval development and fecundity of females *Aedes albopictus* (Skuse). *International Journal of Life Science and Medical Research, 2*(1), 63-67.

Neira, M., Lacroix, R., Cáceres, L., Kaiser, P. E., Young, J., Pineda, L., ... & McKemey, A. (2014). Estimation of Aedes aegypti (Diptera: Culicidae) population size and adult male survival in an urban area in Panama. *Memórias do Instituto Oswaldo Cruz, 109*(7), 879-886.

Rao, B. B., Harikumar, P. S., Jayakrishnan, T., & George, B. (2011). Characteristics of Aedes (Stegomyia) albopictus skuse (Diptera: Culicidae) breeding sites. *Southeast Asian Journal of Tropical Medicine and Public Health, 42*(5), 1077-1081.

Reiskind, M. H., & Zarrabi, A. A. (2012). Is bigger really bigger? Differential responses to temperature in measures of body size of the mosquito, *Aedes albopictus*. *Journal of Insect Physiology, 58*(7), 911-917.

Rowbottom, R., Carver, S., Barmuta, L. A., Weinstein, P., Foo, D., & Allen, G. R. (2015). Resource limitation, controphic ostracod density and larval mosquito development. *PLoS ONE, 10*(11), 1-13.

Rozilawati, H., Zairi, J., & Adanan, C. R. (2017). Seasonal abundance of *Aedes albopictus* in selected urban and suburban areas in Penang, Malaysia. *Tropical Biomedicine, 24*(1), 83-94.

Sairi, F. A. M., Dom, N. C., & Camalxaman, S. N. (2016). Infestation profile of Aedes mosquitoes in multi-storey buildings in Selangor, Malaysia. *Social and Behavioral Sciences, 222*(2016), 283-289.

Schneider, J. R., Morrison, A. C., Astete, H., Scott, T. W., & Wilson, M. L. (2004). Adult size and distribution of Aedes aegypti (Diptera : Culicidae) associated with larval habitats in Iquitos, Peru. *Journal of medical entomology, 41*(4), 634-642.

Teng, H. J., & Apperson, C. H. (2000). Development and survival of immature *Aedes albopictus* and Aedes triseriatus (Diptera: Culicidae) in the laboratory: effects of density, food, and competition on response to temperature. *Entomological Society of America, 37*(1), 40-52.

Vidal, P. O., & Suesdek, L. (2012). Comparison of wing geometry data and genetic data for assessing the population structure of Aedes aegypti. *Infection, Genetics and Evolution, 12*(3), 591-596.

Vijayakumar, K., Kumar, T. S., Nujum, Z. T., Umarul, F., & Kuriakose, A. (2014). A study on container breeding mosquitoes with special reference to Aedes (Stegomyia) aegypti and *Aedes albopictus* in Thiruvananthapuram district, India. *Journal of Vector Borne Diseases, 51*(1), 27-32.

Yoshioka, M., Couret, J., Kim, F., McMillan, J., Burkot, T. R., Dotson, E. M., … & Vazques-Prokopec, G. M. (2012). Diet and density dependent competition affect larval performance and oviposition site selection in the mosquito species Aedes albopictus (Diptera: Culicidae). *Parasites and Vectors, 5*(1), 225-236.

Pertanika J. Sci. & Technol. 27 (4): 1603 - 1614 (2019)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Monitoring of Fracture Healing Process by Acousto-Ultrasonic Technique

Sharad Shrivastava[1]* and Ravi Prakash[2]

[1]*Department of Mechanical Engineering, Birla Institute of Technology and Science, Pilani, Pilani Campus, Vidya Vihar, Pilani, Jhunjhunu Rajasthan India*
[2]*Amity Institute of Technology, Amity University Campus, Sector 125, Noida 201313 (Adjoining New Delhi), India*

ABSTRACT

Presently the radiological examination is widely used for the assessment of *in vivo* bone condition. In some clinical problems such as diagnosis of the point of union of a fracture; the manual assessment is used along with radiological examination. Uncertainty regarding the significance of the radiographic and clinical findings may result in unnecessary long immobilization periods which can produce discomfort for the patient, as well as possible stiffness of the joints and sometimes permanent loss of mobility specially in older population. This paper deals with the *in vivo* analysis of bones by Acousto-Ultrasonic technique. The fracture was created through surgery on one of the limbs of rabbit and healing process was monitored through acousto-ultrasonic technique. A new index known as bone healing index was defined to calculate the end point of healing. In most of the cases, the bone healing index value was found to be 0.80, indicating that 80 % of the strength could only be restored at the completion of healing process.

Keywords: Acousto-ultrasonic, bone healing index, fracture healing, stress wave factor

ARTICLE INFO

Article history:
Received: 30 September 2018
Accepted: 13 May 2019
Published: 21 October 2019

E-mail addresses:
sharad_shrivastava@pilani.bits-pilani.ac.in (Sharad Shrivastava)
raviprakash.ravi@gmail.com (Ravi Prakash)
*Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

INTRODUCTION

Bone fractures are the most common injury treated by the orthopaedic surgeons. Praemer et al. (1992) reported that millions of fractures occurred every year worldwide, with nearly 6.2 million fractures reported annually in the United States alone. Bone fractures can result from either trauma or pathological conditions. Fracture occurs in bones when the normal loading exceeds the load to which the bone has adapted itself during its growth and development and keeping this in mind, one may induce a trauma related fracture in bones. A pathological fracture occurs under normal loading conditions after the bone has been weakened by diseases, such as osteoporosis or bone tumors. Bone fracture healing is a complex and dynamic process of regeneration that slowly and gradually restores the structural integrity and mechanical function of the bone. The healing process involves formation of callus bone followed by change of callus bones into cortical bones and this process may take a few months. Einhorn (1995) study through statistical approach, revealed that in about 5–10% of the fractures that occurred annually, impairment of the healing process might lead to delayed union or non-union requiring further conservative or even surgical procedures.

Monitoring of fracture healing process is done by any suitable process which can evaluate the status of regaining the original strength of the healing bone. The monitoring should also be able to detect any complications in the bone healing process, and accurately assess proper healing. Blokhuis et al. (2001) reported that in routine practice, evaluation of fracture healing was performed by serial clinical and radiographic examinations, both of which depended on the orthopaedic surgeon's expertise and clinical judgment. Development of a quantitative technique would be more helpful for taking decision regarding removal of plaster cast, fixators etc. at an appropriate time.

Several non-invasive techniques have been reported in the literature for proper monitoring of fracture healing process. These include bone densitometry (Augat et al., 1997; Markel et al., 1991; Yoon & Yu, 2018), vibrational analysis (Cunningham et al., 1990; Nakatsuchi et al.,1996; Nikiforidis et al., 1990; Wang et al., 2017), acoustic emission (Aggelis et al., 2015; Claes et al., 2002; Hirasawa et al., 2002; Karaduman et al., 2018 Watanabe et al., 2001) and the attachment of strain gauges to external fixation devices as investigated by Moorcroft et al. (2001), Maffulli and Thornton (1995) and Richardson et al. (1994). All these methods aim to determine the stiffness of the healing bone, either directly or indirectly. Their ability to provide meaningful indications of healing has been validated through ex vivo and in vivo studies. However all these techniques are usually affected by extrinsic bone properties, such as bone gross geometry, and fracture type. In addition, these techniques require appropriate clinical setting, necessitating the intervention of a specialist to configure the set-up, and a number of them require removal of external fixators during examination. Ultrasonic methods have also been used for the monitoring of bone fracture healing. Although some researchers have employed ultrasonography (Moed et al., 1998), and power Doppler ultrasonography (Risselada et al., 2006), to assess the appearance and neo-vascularisation of the callus tissue during healing, the majority of the studies have utilized quantitative ultrasound techniques. The axial transmission approach offers a unique way to examine fractures in long bones, such as tibia and radius. In axial transmission approach, a transmitter and a receiver are placed in direct contact with the skin on either side of the fracture, as is illustrated in Figure 1. The emitted ultrasonic waves propagate from the transmitter to the receiver along the long axis of the bone. Due to the complex structure of callus tissues and cortical bone, an appreciable change in the ultrasound velocity across the fracture occurs when compared to measurement of ultrasonic velocity in intact bone. Similar changes have been observed with regard to other propagation characteristics, such as ultrasonic attenuation parameter and dispersion characteristics.

The application of quantitative ultrasound to monitor the healing process has been investigated through animal experimentations and also in clinical studies over a long period of time. Simple experiments on bone phantoms and bone specimens have also been performed, aiming to examine the effect of the fracture characteristics (gap width and depth) on the measured quantities. The recent introduction of computational methods into bone research has extended our understanding of the underlying propagation phenomena and has helped researchers to propose new measurement techniques, such as the use of guided waves.

By considering the fact that two of the main functions of long bones are to support the body weight and locomotion, it is essential that the healing of a fracture should be assessed mainly in terms of increasing mechanical strength.

In this work an attempt has been made to use a combination of acoustic emission and ultrasonic technique for monitoring the fracture healing process. This is the first study wherein acousto-ultrasonic technique has been used for monitoring fracture healing process. Animal experimentation with rabbits were conducted using acousto-ultrasonic technique for monitoring the fracture healing process.

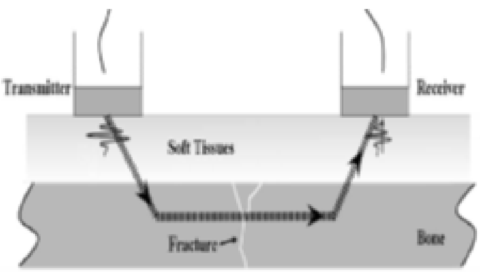


Figure 1. The axial transmission technique for ultrasonic evaluation of fracture healing process in long bones.

EXPERIMENTAL STUDY

An experimental work with rabbits was conducted in the Animal House of Birla Institute of Technology and Science (BITS) Pilani. Twenty (20) healthy rabbits of both sexes were selected from the groups available. The average weights of the rabbits were ranging from 1.5 kg to 2.0 kg and the approximate age of the rabbits were 24 weeks. They were fed with standard diet during the whole experimentation period. The whole experimental work was approved by Animal Ethics Committee (Ref No.: IAEC/RES/12/03).

Creating Closed Fracture

Many investigators used several animal fracture models for the study of fracture healing, even with inherent advantages and disadvantages. The adjustments have been made between the reproducibility of osteotomy and the actual fracture. Generally, the production of real fractures increases the chances of variation in fracture site and location, which in turn results in making retesting more difficult. As it is difficult to grip the whole bone in vivo; the most common fracture process adopted by the researchers is bending (Ashhurst et al., 1982) and (Claes & Cunningham, 2009), where the bone can be supported against two points and a load can be applied from the opposite side. When a single point loading is adopted, the location of the fracture is determined by the loading point, and the mode is termed as "three point bending". A more uniform stress distribution over the test section is obtained when using two parallel loading points within the test span, termed as "four point bending", but the exact location and direction of the fracture is not so well controlled. Burstein and Frankel (1971) reported that the bone fractures at the weakest section as it would be expected in normal service conditions. They proposed a device to create a closed fracture in rats. However, all the methods mentioned above required very sophisticated equipment, and expert post-treatment care for proper healing. The experimental fracture created sometimes separated the rabbit's tibia in two parts, which needs intramedullary nailing to keep them aligned. This again required skill of a surgeon. Keeping in view all of the limitations; we went for creating a surgical fracture giving an oblique cut of 2 to 2.5mm depth in rabbits' left tibia. As this is only the first attempt of its kind, the main goal is to investigate the application of acousto-ultrasonic technique for monitoring the process of fracture healing. The type of fracture produced is not important at this point of study, but definitely it will be needed in the future study once the proof of the acousto-ultrasonic technique has been obtained.

Procedure for Creating Surgical Fracture

Pre-Requirements. The rabbits were kept fasting for nearly 6 hours to avoid any complication due to anaesthesia. The rabbits were anesthetized using a combination of xylazin (5-6 mg/kg body weight) and ketamine hydrochloride (9-13 mg/kg body weight)). Before starting the surgery, all surgical equipments, cottons, distilled water, surgery gown, green cotton cloth were autoclaved at 121°C, for 20 minutes.

Procedure. The Fibula is thin and there is an elongated interosseous space between the two bones. Comparatively less musculature is present on the medial aspect of bone. This was the reason for choosing rabbits' tibia for this study. The following procedure is adapted for creating surgical fracture in rabbit's tibia:

1) Legs of rabbit were washed with disinfectant soap (savlon) water and with clean cloth.

2) The hairs on the right hind legs were removed with the help of scissors and commercial hair remover.

3) The skin was then cleaned with 70% ethyl alcohol and povidon was applied as an anti-bacterial reagent.

4) Local anesthesia lignocaine hydrochloride (approx. 2ml) was in filtered at different places on the medial aspect of tibia region.

5) The skin was removed by giving a vertical incision of 40-50 mm with the help of surgical blade.

6) Facia were removed with the help of scalpel.

7) The muscles were removed by giving irregular cut with the help of point to blunt scissor.

8) At the distal end of posterior side, deep digital flexor muscle was separated from tibia by blunt incisions, at proximal third, bone was cut to the depth of 2.5 - 3mm by electric saw.

9) The muscles were sutured by catgut no.2 in lock and stitch method.

10) The facia and skin was sutured by silk no.2 by horizontal mattress method.

Figure 2 shows the surgical fracture created for the study.

Post Surgical Treatment. The post-surgical treatment provided to the rabbits is as follows:

1) The rabbits hind leg was immobilized by applying Plaster of Paris cast from the stifle joint to hock joint, immobilizing the ankle joint in zero position (neutral) and the knee joint in 900 of flexion.

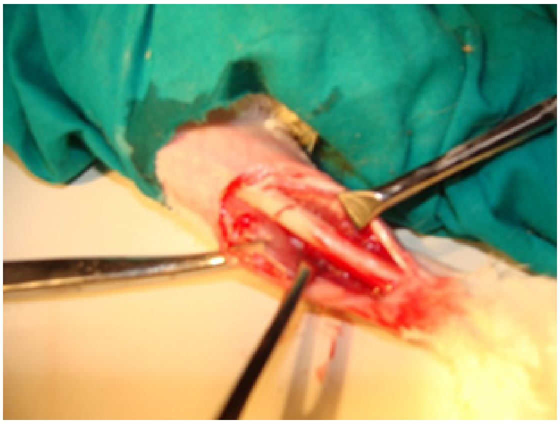


Figure 2. Photograph of surgical fracture in right limb

2) The size of the plaster cast was reduced 3 weeks post operatively, as it was cut just distal to the knee and proximal to the ankles permitting full mobility of these joints.

3) The very next day after the surgery 3-4 holes were drilled on the plaster of Paris near the surgical site, to provide air passage.

4) The antibiotic/painkiller injection fortivir (0.5 ml I/M) and

Diclab (0.30 ml I/M) were given to all the rabbits for 3-4 days.

5)The injection T.T (0.25ml I/M) was also given to all the rabbits for pain relaxation. The antibiotic injection was given continuously for three days.

6) Few rabbits, which were given injection Belamyl (0.5 ml I/M) and injection cobacal D (0.75 ml I/M), showed symptoms of anorexia.

All the rabbits were kept on healthy diet during the study.

ACOUSTO-ULTRASONIC MEASUREMENT

Acousto-ultrasonic measurements were made from the healing limb and also from the contra lateral limb at every two week during the healing period. The contra lateral limb was used as a control for the acousto-ultrasonic measurements. The plaster cast was carefully removed during the acousto-ultrasonic measurements and if required applied again afterwards. In most of the cases the plaster cast was totally removed after 8 weeks. As the fracture was just a cut of 2-2.5 mm only, by examining the X-ray after 8 weeks, the decision was taken that whether the plaster cast should be applied again or not. Figure 3 shows the arrangement for acousto-ultrasonic measurement.

The pocket hand held acousto-ultrasonic unit from NDT Automation, a member of MISTRAS group has two spring-loaded, wheeled, rolling transducers attached with it. One sensor is the ultrasonic pulser and the other sensor is the acoustic emission receiver. The distance between the two sensors is one inch fixed. Here using the ultrasonic pulser, the low amplitude ultrasonic pulses of fixed frequency 250 kHz were injected in to the rabbit's tibia for the duration of 100µs. The pulser has the capability of generating the burst frequencies in the range of 50 kHz-1MHz, and the output voltage of 20 Volts peak to peak. The 25% of the noise signal were removed through the inbuilt system within the unit. The calibration of the amplitude height, threshold and other equipment settings were done with respect to contralateral limb before testing the healing bone. Once the calibration was done, the entire equipment setting was kept constant during the whole monitoring process for all the rabbits. The threshold setting was done with the help of setting detection gate facility with the instruments. The stress waves after interacting within the tibia bone were picked up by the 10kHz – 1MHz frequency response acoustic emission sensor. The care was taken while placing the sensors over the bone while taking the measurement. The instrument was somewhat modified, so that adequate amount of pressure is maintained so that proper sitting of the sensors could be assured. The sensor placement sites were so chosen such that the soft tissue thickness is the minimum at those places. The site chosen for the surgical fracture had the minimum soft tissue thickness. The received signal was them amplified by the built-in preamplifier of 40dB gain and bandwidth of 10kHz-1MHz. The equipment had the facility of activating the high pass filter and low pass filter within it. The high pass filter frequency of 100 kHz and the frequency of 400 kHz for low pass filter were chosen for the in vivo study.

The Pocket AU system has its own built-in AU software program that uses A-Scan and C-scan analysis to perform Acousto-ultrasonic inspection. Only the A-Scan analysis feature was used for this study, as the scanning of the rabbit's tibia was not possible. The received signal was digitised through the dedicated feature extraction processor within the unit. The waveform parameters were then exported and converted to ASCII files and transferred to the personal computer for further signal processing.

The stress wave factor for the study is taken as the energy content of AU signals. The relative energy is given by formula (1) [12]

$$\text{Energy integral} = \text{SWF} = \int_{t1}^{t2} [v(t)]^2 dt \qquad 1$$

Where, v(t) is the amplitude distribution with respect to time.

The Stress Wave factor (energy content) corresponding to both the control and healing limb were measured after every two weeks during the healing period up to 12 weeks from the date the fracture was made. A Healing Index which is a ratio of SWF of healing limb and SWF of control limb was calculated.

Radiographic examination was also made every month for qualitative assessment. The software was developed in MATLAB 6.05 for calculating the different parameters of the acoustic emission signal and for calculating the healing index, the term used for monitoring the healing process,

RESULTS AND DISCUSSIONS

The software specially developed for the present study took the AU waveforms recorded during the AU measurements as input and process the signals to find different parameters such as energy content of the signal, peak amplitude, rise time, and time of flight. The signals were also plotted directly on the screen and the calculated Healing index was also displayed.

Figure 3 gives the details of the acousto-ultrasonic measurements made in rabbit No. 4 just after the surgery. It clearly mentions the various parameters calculated for both the control and healing limb just after the fracture was created. In Figure 3 output for "good" stands for control limb and "bad" for Healing limb.

Figure 4 displays the energy content (used as SWF for the present study) for both the control and healing limbs of rabbit no. 4 as well as healing index just after the surgery. The healing index was found to be 0.003, as the maximum signal experienced attenuation due to the presence of fracture. The X-ray picture of the limb post-surgery is shown in Figure 5.

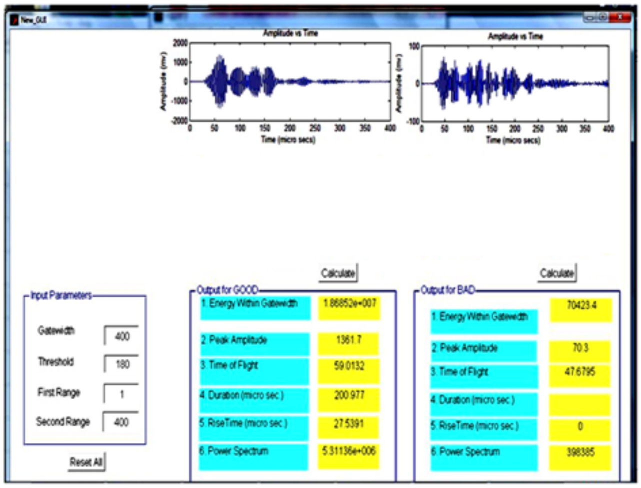


Figure 3. Acousto-ultrasonic parameters calculated for rabbit number 4 just after surgery

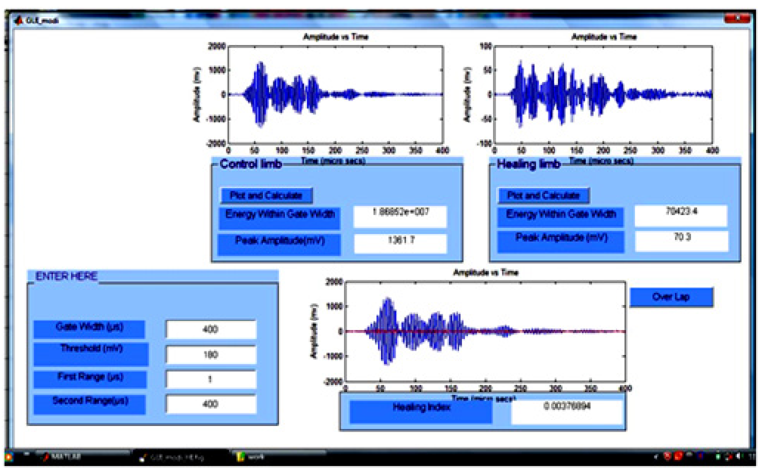


Figure 4. Acousto-ultrasonic measurements for rabbit No. 4 just after the surgery

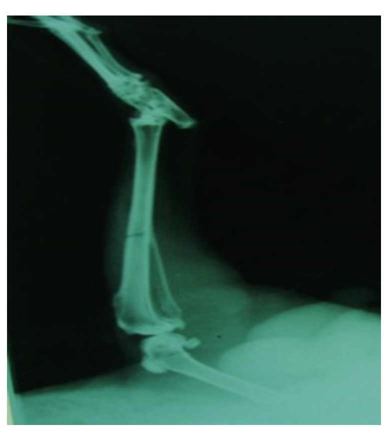


Figure 5. X-radiograph of rabbit number 4 just post-surgery

The value of bone healing index in the fourth week during the healing period indicates that the callus formation around the fracture has taken place at this time. The value of the healing index clearly shows an increasing trend during the healing period. The value of the healing index after week 12 was found to nearly constant when measured in the fourteenth week. Hence the healing index value of 0.86 obtained after week twelfth was considered to be the end point of healing process.

Figure 6 and 7 show the value of the healing index and the X-ray photograph of rabbit 4 at the end of week twelfth. The value of the healing index clearly indicates that it increased considerably till week eighth but after that its value increased very slowly. This was due to slow remodeling process in the final stage of fracture healing process, which continued after week 12 also.

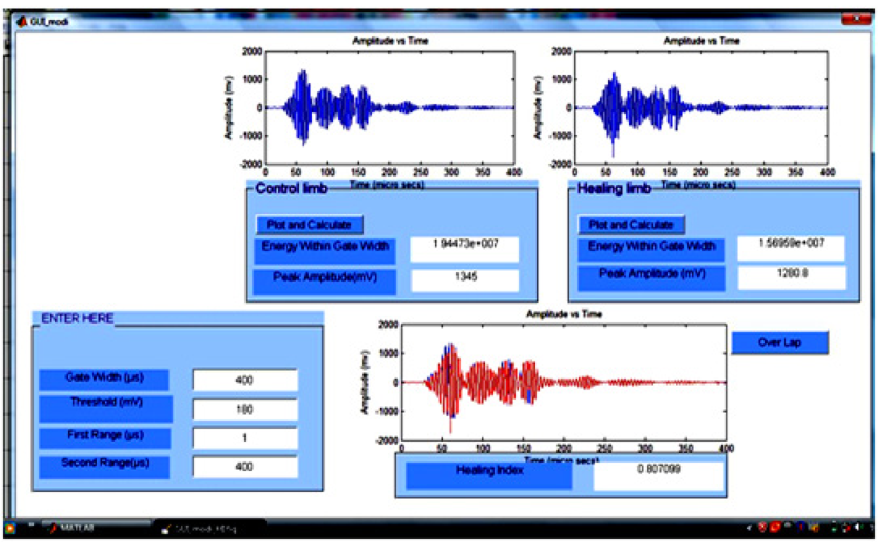


Figure 6. Acousto-ultrasonic measurements for rabbit number 4 at twelfth week

Figure 7. X-ray of rabbit number 4 at twelfth week

Subsequent graphs Figure 8 to 10, shows the Bone healing index values for rabbits number 10, 11 and 13 respectively.

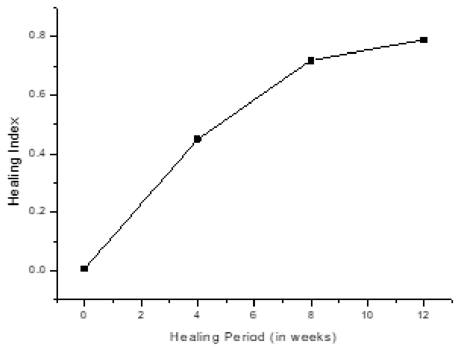


Figure 8. Healing Index values for rabbit number 10 during the healing period

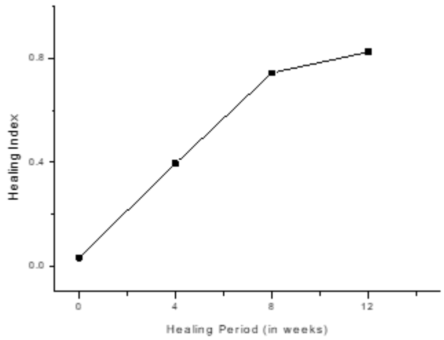


Figure 9. Healing Index values for rabbit number 11 during the healing period

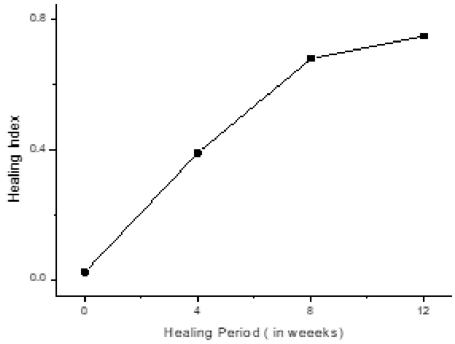


Figure10. Healing Index values for rabbit number 13 during the healing period

CONCLUSION

The results of the present experimental study indicate that Acousto-ultrasonic technique can be used for monitoring the fracture healing process.

The calculated Healing Index from the acousto-ultrasonic measurements was found to be increasing uniformly during the healing period. The end point of the healing process was indicated by the more or less constant value of Healing Index value at 12-14 weeks of the healing period.

The X-ray examination clearly show the healing has almost completed after the eighth week, but the Healing Index value at the same time does not confirm it. Hence it can be concluded that the healing index value which was calculated with respect to energy content of the signal, clearly assembled the strength of the bone also. Thus, it is concluded that the acousto-ultrasonic technique, apart from monitoring the fracture healing process also gives the information of the strength of the bone achieved. The correlation between the energy content (SWF) and the strength of the bone has already been validated in the *in vitro* study with this technique. Due to the limitations we could not perform the same correlation for the *in vivo* studies.

ACKNOWLEDGEMENT

The authors wish to acknowledge Dr. Sushil Yadav and his team for his help in creating the experimental fracture. Financial assistance provided by Department of Science and Technology, Government of India is also duly acknowledged.

REFERENCES

Aggelis, D., Strantza, M., Louis, O., Boulpaep, F., Polyzos, D., & van Hemelrijck, D. (2015). Fracture of human femur tissue monitored by acoustic emission sensors. *Sensors, 15*(3), 5803-5819.

Ashhurst, D. E., Hogg, J., & Perren, S. M. (1982). A method for making reproducible experimental fractures of the rabbit tibia. *Injury, 14*(3), 236-242.

Augat, P., Merk, J., Genant, H. K., & Claes, L. (1997). Quantitative assessment of experimental fracture repair by peripheral computed tomography. *Calcified Tissue International, 60*(2), 194-199.

Blokhuis, T. J., De Bruine, J. H. D., Bramer, J. A. M., Den Boer, F. C., Bakker, F. C., Patka, P., ... & Manoliu, R. A. (2001). The reliability of plain radiography in experimental fracture healing. *Skeletal Radiology, 30*(3), 151-156.

Burstein, A. H., & Frankel, V. H. (1971). A standard test for laboratory animal bone. *Journal of Biomechanics, 4*(2), 155-158.

Claes, L. E., & Cunningham, J. L. (2009). Monitoring the mechanical properties of healing bone. Clinical Orthopaedics and Related Research, 467(8), 1964-1971.

Claes, L., Grass, R., Schmickal, T., Kisse, B., Eggers, C., Gerngross, H., ... & Wentzensen, A. (2002). Monitoring and healing analysis of 100 tibial shaft fractures. *Langenbeck's Archives of Surgery, 387*(3-4), 146-152.

Cunningham, J. L., Kenwright, J., & Kershaw, C. J. (1990). Biomechanical measurement of fracture healing. *Journal of Medical Engineering and Technology, 14*(3), 92-101.

Einhorn, T. A. (1995). Enhancement of fracture-healing. *The Journal of Bone and Joint Surgery, 77*(6), 940-956.

Hirasawa, Y., Takai, S., Kim, W. C., Takenaka, N., Yoshino, N., & Watanabe, Y. (2002). Biomechanical monitoring of healing bone based on acoustic emission technology. *Clinical Orthopaedics and Related Research, 402,* 236-244.

Karaduman, D., Bircan, D. A., & Cetin, A. (2018). Assessment of Crack Initiation and Propagation in Bone Using Acoustic Emission (AE) Techniques. *Journal of Mechanics in Medicine and Biology, 18*(03), 1850031.

Maffulli, N., & Thornton, A. (1995). Ultrasonographic appearance of external callus in long-bone fractures. *Injury, 26*(1), 5-12.

Markel, M. D., Morin, R. L., Wikenheiser, M. A., Lewallen, D. G., & Chao, E. Y. (1991). Quantitative CT for the evaluation of bone healing. *Calcified Tissue International, 49*(6), 427-432.

Moed, B. R., Subramanian, S., van Holsbeeck, M., Watson, J. T., Cramer, K. E., Karges, D. E., ... & Bouffard, J. A. (1998). Ultrasound for the early diagnosis of tibial fracture healing after static interlocked nailing without reaming: clinical results. *Journal of Orthopaedic Trauma, 12*(3), 206-213.

Moorcroft, C. I., Ogrodnik, P. J., Thomas, P. B., & Wade, R. H. (2001). Mechanical properties of callus in human tibial fractures: a preliminary investigation. *Clinical Biomechanics, 16*(9), 776-782.

Nakatsuchi, Y., Tsuchikane, A., & Nomura, A. (1996). Assessment of fracture healing in the tibia using the impulse response method. *Journal of Orthopaedic Trauma, 10*(1), 50-62.

Nikiforidis, G., Bezerianos, A., Dimarogonas, A., & Sutherland, C. (1990). Monitoring of fracture healing by lateral and axial vibration analysis. *Journal of Biomechanics, 23*(4), 323-330.

Praemer, A., Furner, S., & Rice, D. P. (1992). *Musculoskeletal conditions in the United States.* Park Ridge, IL: In American Academy of Orthopedic Surgeons.

Richardson, J. B., Cunningham, J. L., Goodship, A. E., O'connor, B. T., & Kenwright, J. (1994). Measuring stiffness can define healing of tibial fractures. *The Journal of Bone and Joint Surgery. British volume, 76*(3), 389-394.

Risselada, M., van Bree, H., Kramer, M., Chiers, K., Duchateau, L., Verleyen, P., & Saunders, J. H. (2006). Evaluation of nonunion fractures in dogs by use of B-mode ultrasonography, power Doppler ultrasonography, radiography, and histologic examination. *American Journal of Veterinary Research, 67*(8), 1354-1361.

Wang, J., Leung, K. S., Chow, S., & Cheung, W. H. (2017). The effect of whole body vibration on fracture healing–a systematic review. *European Cells and Materials, 34,* 108-127.

Watanabe, Y., Takai, S., Arai, Y., Yoshino, N., & Hirasawa, Y. (2001). Prediction of mechanical properties of healing fractures using acoustic emission. *Journal of Orthopaedic Research, 19*(4), 548-553.

Yoon, B. H., & Yu, W. (2018). Clinical utility of biochemical marker of bone turnover: fracture risk prediction and bone healing. *Journal of Bone Metabolism, 25*(2), 73-78.

Pertanika J. Sci. & Technol. 27 (4): 1615 - 1624 (2019)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

MRN Complex and ATM Kinase Inhibitors Impacts towards UVC-Treated Zebrafish Embryonic Development

Shazrul Fazry[1,2]*, Malina Kumaran[1], Nahid Khalili[1], Ibrahim Mahmood[1], Herryawan Ryadi Eziwar Dyari[1,2], Nur Hidayah Jamar[1,2], Douglas Law[3], Nik Marzuki Sidik[3], Nor Azian Abdul Murad[4] and Yi Chieh Lim[5]

[1]*Faculty of Science and Technology, Universiti Kebangsaan Malaysia, 43600 UKM, Bangi, Selangor, Malaysia*
[2]*Chini Lake Research Center, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, 43600 UKM, Bangi, Selangor, Malaysia*
[3]*Faculty of Agro Based Industry, Universiti Malaysia Kelantan, Jeli Campus, Locked Bag 100, 17600 UMK, Jeli, Kelantan, Malaysia*
[4]*UKM Medical Molecular Biology Institute, UKM Medical Center, Jalan Ya'acob Latiff, Bandar Tun Razak, 56000 Cheras, Kuala Lumpur, Malaysia*
[5]*dkDanish Cancer Society, Research Centre Strand Boulevard, 49DK-2100 Copenhagen, Denmark*

ABSTRACT

The zebrafish (*Danio rerio*) has become a prevailing vertebrate model for developmental biology studies due to its ease of care, rapid embryogenesis stages development and translucent embryos. In this studies, ATM Kinase and MRN complex role as DNA damage response proteins during embryogenesis was examined by using specific MRN complex (Mirin) and ATM Kinase inhibitors (Ku60019 and Ku55933). To create DNA lesions in zebrafish, embryos at mid-blastula transition (MBT) stage were exposed to inhibitors (Mirin, Ku60019 or Ku55933) and later exposed to UVC irradiation wavelength of between 100 to 280 nm. Hatching but with visible physical deformation was observed for embryos treated with Mirin, Ku60019 or Ku55933 and UVC exposure at concentration of 3μM, 1.5 nM and 3nM or lower, respectively up to 72 hours-post fertilisation (hpf). On the other hand, no deformities were observed for all control as well as mock treated embryos. This study confirmed that DNA damage response proteins are crucial during embryo development to prevent undesired abnormal

ARTICLE INFO

Article history:
Received: 24 April 2019
Accepted: 02 July 2019
Published: 21 October 2019

E-mail addresses:
shazrul@ukm.edu.my (Shazrul Fazry)
malina.kumaran@gmail.com (Malina Kumaran)
nahidkhalili599@gmail.com (Nahid Khalili)
ibrahim_m_mahdy@yahoo.com (Ibrahim Mahmood)
herry@ukm.edu.my (Herryawan Ryadi Eziwar Dyari)
njamar@ukm.edu.my (Nur Hidayah Jamar)
douglas.law@gmail.com (Douglas Law)
nikmarzuki@UMK.edu.my (Nik Marzuki Sidik)
nor_azian@ppukm.ukm.edu.my (Nor Azian Abdul Murad)
yilim@cancer.dk (Yi Chieh Lim)
* Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

biological development. Thus, it is proven that protein inhibitors are a cheaper alternative in valuating specific protein roles during embryogenesis compared to both genomic and transcription modification tools.

Keywords: ATM Kinase, *Danio rerio*, DNA damage, MRN complex, zebrafish

INTRODUCTION

DNA lesion can occur consequently from cell exposure to endogenous and exogenous genotoxic pressure. Aside from that, DNA also prone lesion during the polymerisation of single strand DNA due to collapse of the replication forks (Lavin, 2007). The MRE11-RAD50-NBN (MRN) complex which consists of meiotic recombination 11 (MRE11), DNA repair protein RAD50 (RAD50) and nibrin (NBN) has been identified to play important role in distinguishing the DNA double strand breaks and activates DNA repair proteins mechanism through ataxia telangiectasia mutated (ATM) kinase (Gatei et al., 2011; Lavin, 2007). Null mutation of either one of the MRN complex proteins disrupt proper embryo development which is lethal (Luo, 1999; Williams et al., 2002; Zhu et al., 2001), signified the crucial function of every single MRN components during this process. The RAD50 gene significance in sustaining genomic stability was recorded through a female patient having a RAD50 deficiency (Waltes et al., 2009). The RAD50 deficiency disorder was later renamed as Nijmegen Breakage Syndrome-like Disorder (DBSLD). DBSLD patients is having similar characterisation usually observed in NBS patients such as short stature, bird-like face, microcephaly and mental retardation (Waltes et al., 2009).

The importance of RAD50 and the implications of any of its mutations studies have largely been *in vitro*. It is important to note that *in vitro* study does not give a clear picture of how disruption of RAD50 function could affect the phenotype of an organism. Some *in vitro* studies even fail to replicate the precise cellular conditions of a mutated organism (Bender et al., 2002). Although previous researches using mice could shed some light on the implications of RAD50 deficiency on the function of the MRN complex, there are many more aspects that remain elusive. Bender et al. (2002) had shown that the RAD50s/s mutant mice exhibited growth defects and cancer predisposition and the outcomes of the RAD50s/s phenotype were not associated with overt defects in the MRN complex. Indicating that subtle alteration of the MRN complex might have deleterious effects. However, the actual importance of RAD50 have yet to be fully understood. Knowing how important the correct expression of RAD50 is in ensuring MRN complex proper functioning leads to another question of how important is RAD50 in ensuring genomic stability.

While studies on RAD50 have been established in mice previously, mainly due to their evolutionary proximity and homology to human, zebrafish (*Danio rerio*) are yet to be utilised in RAD50-related researches. Zebrafish can be used to explore the genetic and physiological mechanisms involved in the functioning of various systems (Meyers, 2018). Zebrafish embryos are particularly an attractive model for *in vivo* studies to their optical transparency and rapid embryonic development; making it possible to observe embryogenesis and organogenesis in real time (Bladen et al., 2005). This can be exploited to observe the effects of the inhibitors on their development and bodily functions (Dale et al., 2009; Driever et al., 1996; Golling et al., 2002; Segner 2009). The aim of this study was to investigate the effects of MRN complex, ATM and ATR Kinase inhibitors on UVC-treated zebrafish embryogenic development.

METHODS

Zebrafish Husbandry

Adult wild type zebrafish were kept at standard condition of 14:10 hour light-dark photoperiod. The water pH and temperature were maintained at 7.4 and 28.5°C, respectively. Eggs produced were screened for fertilization under light microscope 1 h post fertilization (hpf). Selected fertilized eggs was incubated in E3 solution (5 mM NaCl, 0.17 mM KCl, 0.33 mM CaCl2, 0.33 mM MgSO4). Animal care and all experimentation were conducted as previously reported by Fazry et al., (2018) and was approved by the UKM ethics committee (Animal ethic approval number: (FST/2015/SHAZRUL/25-MAR./670-MAR.-2015-DEC.-2016).

Treatment with Inhibitors

Selected single cell zebrafish embryos at mid-blastula transition mid-blastula transition (MBT) stage were treated individually with 200 µL using Mirin (MRN complex inhibitor) (Sigma-Aldrich, USA), Ku60019 (ATM Kinase inhibitor)(Sigma-Aldrich, USA) and Ku55933 (ATM and ATR Kinase inhibitor)(Sigma-Aldrich, USA) prior to UVC irradiation at room temperature. Inhibitor concentrations used in the treatment were 3, 6 and 12 µM for Mirin; 1.5, 3 and 6 nM for Ku60019; and 3, 6 and 12 nM for Ku55933. The concentration of inhibitors used were based on the half maximal inhibitory concentration (IC50) of the inhibitors (Dupré et al., 2008; Negi & Brown, 2015; Williams et al., 2013).

Irradiating the Embryos with UVC

The embryos that were exposed to the inhibitors were then incubated at 28.5°C for 1 hour prior to UVC radiation to ensure that DNA damage indeed occurs. UVC has a wavelength of 100-280 nM which represents the most lethal form of radiation in comparison to UVB and UVA. UVC is utilized to ensure maximum exposure of UV ray that will induce DNA damage in the developing embryos (Clemente et al., 2014). Embryos were exposed to radiation using a UVC light (4.56 watt/m^2) in a customized box. A dosage of 800 Joules/m^2 was given to the inhibitor-treated embryos by manipulating the duration of the UVC exposure (1 watt/m^2 = 1 Joules/second/m^2) (Ledo, 1993; Kaur et al., 2018). The non-UVC treated embryos were maintained under standard laboratory lighting. The embryos were left to be incubated at 28.5°C and their survival and development were evaluated at 24, 48 and 72 hpf, under light microscope at 4× magnification.

Statistical Analysis

Values of experimental results shown in figures were the mean of at least three determinations (±standard deviation).

RESULTS AND DISCUSSIONS

In this study, we observe significant changes in the phenotypes of inhibitor and UVC-treated embryos at 3 days' post fertilisation (dpf) in comparison with control embryos. The data collected were distributed into analysing the survivability of the embryos and the characterisation of deformities observed in embryos that had survived the treatment at the end of 3 dpf. A total of 15 embryos were used in triplicates of 5 embryos (Figure 1).

Zebrafish embryos that were exposed to inhibitors within the MBT phase without prior exposure to UVC radiation had a good survivability (Figure 1A). All the embryos treated with different concentration of Mirin had 100% survivability. The ATM kinase inhibitor Ku60019-treated embryos had slightly lower survivability than Mirin, with the 1.5 nM displaying the lowest survivability at 86%. The Ku55933 treated embryos had the least survivability with 80% of embryo alive at the lowest concentration 3 nM, and 86% for both 6 and 12 nM treated embryos. This result indicates that the Ku55933 is a more potent inhibitor in comparison to Mirin and Ku60019 as the lowest concentration of 3 nM can result in the lowest survivability of embryos. This may be contributed by the mechanism at which Ku55933 works since it inhibits both the ATM and ATR kinase activity (Chwastek et al., 2017). Hence, it can halt the downstream signalling of the DNA damage response more effectively than Mirin and Ku60019.

In the UVC treated embryos, the survivability dropped significantly as observed in the control embryos (Figure 1B). The survivability of Mirin-treated embryos was the highest at 53% when the lowest concentration of 3 µM is used. The lowest survivability was observed at 20% when the embryos were treated with 6 µM Mirin. In the Ku60019 treated embryos, the survivability drops from 46% to 40% as the concentration increases from 1.5 to 6 nM, respectively. Embryos treated with 12 nM Ku55933 had the lowest survivability of only 26%. Based on this data, we can conclude that the survival of inhibitor treated embryos decreases when exposed to UVC. However, percentage of survived inhibitors treated embryos did not portray the actual effect of UVC on MRN complex inhibited zebrafish embryos. This was due to the nature of random pyrimidine photodimers created by UV irradiation (Mullenders, 2018; Nair & Loppnow, 2019; Sugiyama & Chen, 2019), which affected the survivability rate of the embryos.

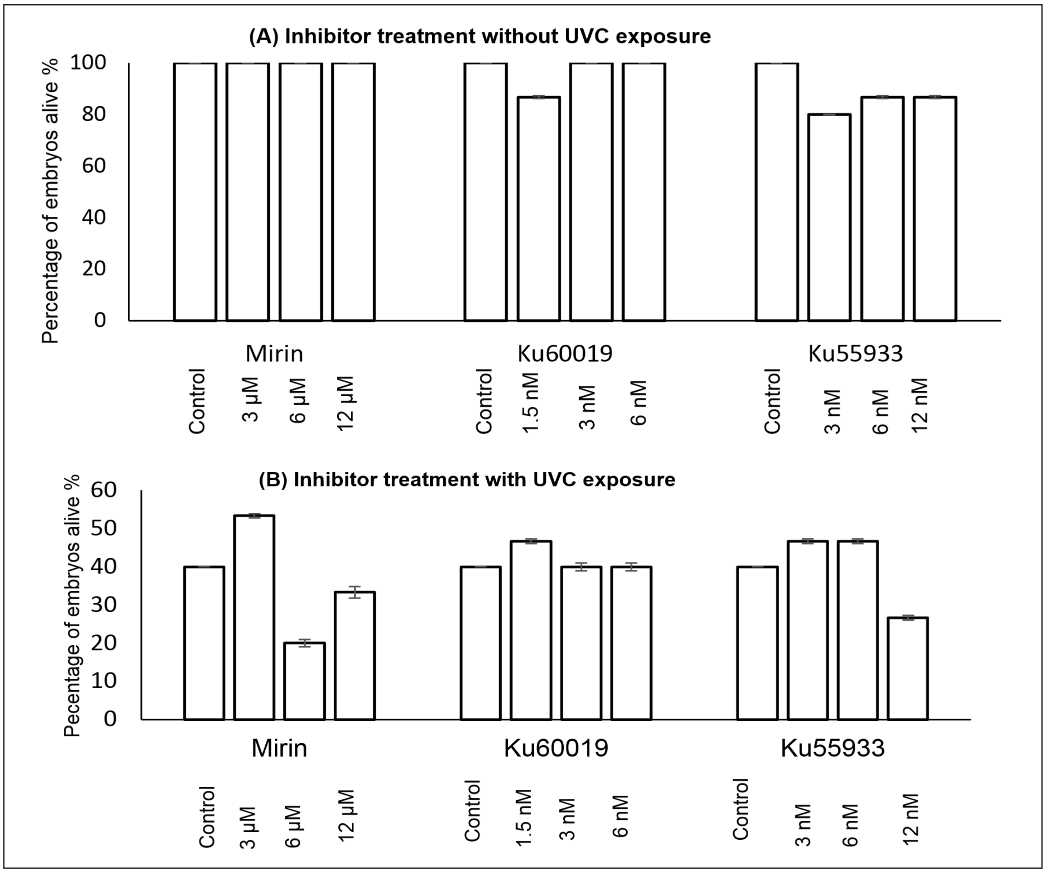


Figure 1. Survivability of embryos with and without UVC exposure at the end of 72 hpf. (A) Graph shows the survivability of embryos without UVC exposure. Embryos were treated with inhibitors at different range of concentration within the MBT stage and were incubated at 28.5°C under standard laboratory lighting. (B) Graph shows the survivability of embryos with UVC exposure. Briefly, the inhibitor treated embryos were incubated for 1 hour at room temperature before being irradiated with UVC. The embryonic development was evaluated at 24, 48 and 72 hpf. Experiments were performed in triplicate and the data are expressed as mean ± SD

Embryos which were alive by the end of 72 hpf were evaluated for the manifestation of any morphological abnormalities (Figure 2). On the other hand, embryos which displayed any deformed phenotype were tabulated as having abnormal development. In this study, only the embryos treated with UVC developed deformities while the non-UVC treated embryos developed normally with or without protein inhibitors (data not shown). The control embryos which were treated with 0.1% DMSO have the lowest rate of deformed embryos, excluding the embryos treated with 6 μM Mirin. The rest of the embryos treated with varying concentration of the inhibitors developed into phenotypically deformed embryos by 72 hpf, with embryos treated with Mirin (3 μM) having the highest rate of deformed embryos at 53%.

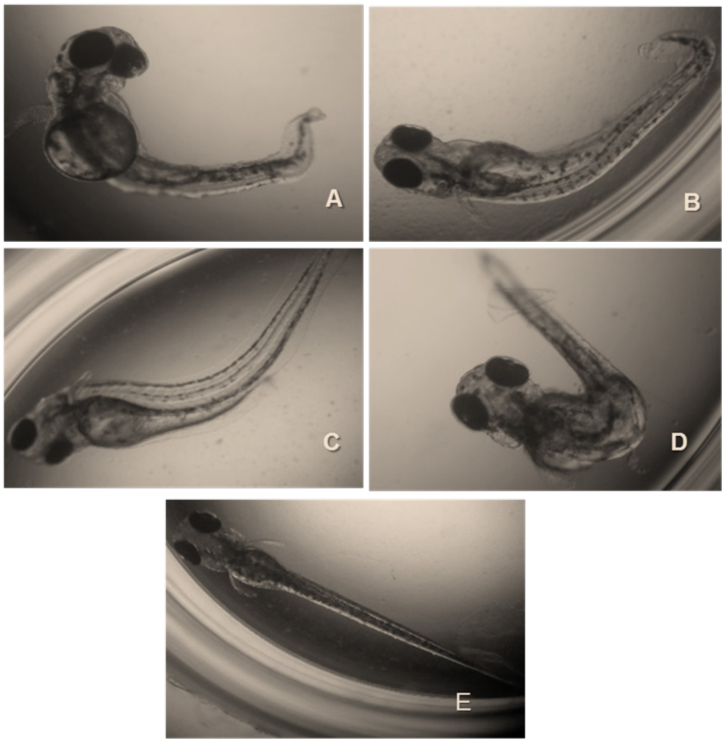


Figure 2. The figure illustrates the deformed phenotypes and the normal phenotypes observed at 72 hpf. (A) Enlarged yolk sac (B) Bent caudal fin (C) Curved-body (D) Bent floor plate and (E) Normal. *The data represents the % of deformed embryos from total embryos which are still alive at 72 hpf.

The majority of embryos treated with UVC developed enlarged yolk sac by the end of 72 hpf. A hundred (100) percent of alive control embryos had enlarged yolk sac, similarly with the rest of the Mirin, Ku60019 and Ku55933 inhibitors with varying concentration (Figure 3A). Interestingly, treatment with 3 µM Mirin and 3 nM Ku60019 showed a decrease of 12.5% and 28.6% in the formation of enlarged yolk sac, respectively. Based on this result, we can assume that the DNA damage inflicted by the UVC affects the development of the yolk sac. The damage is prominent enough to cause the phenotype to have almost 100% manifestation across the range of inhibitors at varying concentrations used.

In the UVC treated embryos, 75% of embryos alive developed bent caudal fin phenotype, indicating that the development of zebrafish embryos caudal tail during embryogenesis were sensitive to DNA damage (Figure 3B). Unexpectedly, the formation of the bent caudal tail phenotype decreased with the additional of Mirin, Ku60019 and Ku55933. This may be due to the inhibition of apoptosis of the damage cells when treated with these inhibitors. Previous studies have reported that both MRN and ATM induced apoptosis on DNA damaged cells (Chwastek et al., 2017, Dubois & Gerber, 2016). We speculate that cells with compromised DNA structure in UVC treated embryos (without

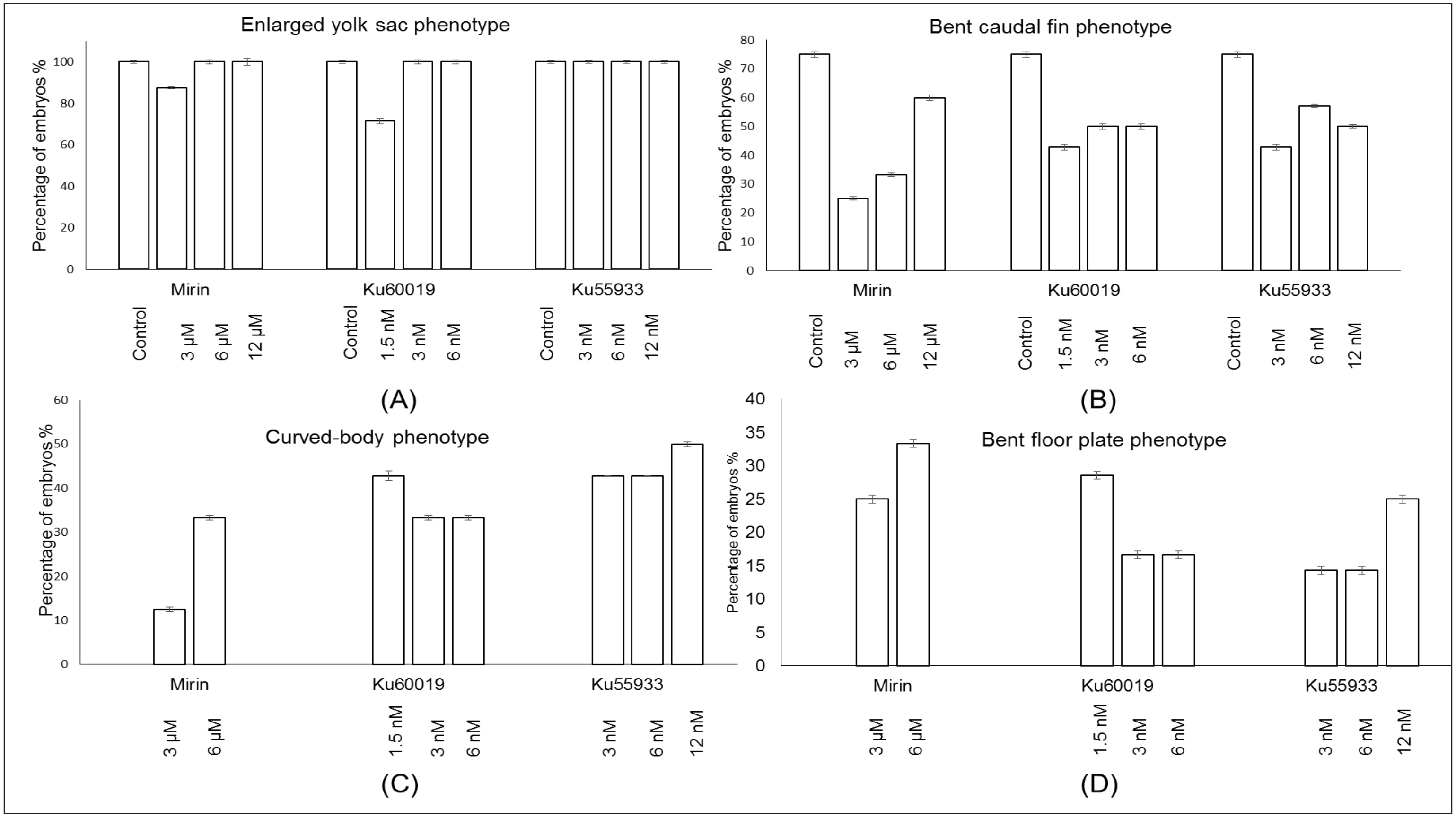


Figure 3. Percentage of embryos that develop deformed phenotypes. These phenotypes are observed only in the UVC treated embryos. (A) Graph illustrates the percentage of embryos that develop enlarged yolk sac by 72 hpf. (B) Graph shows the rate of embryos which have develop bent caudal fin. (C) The manifestation of curved-body phenotype among the alive embryos is shown across the inhibitor type and concentrations. (D) Graph shows the % of embryos that develop bent floor plate phenotype. Alive and morphologically deformed embryos at 72 hpf were evaluated for characteristic differences in the type of deformities observed. The deformities were then categorised into these 4 deformed phenotypes. Experiments were performed in triplicate. Error bar showed the standard error of the mean. Asterisk (*) showed the significant ($p < 0.05$) difference between treated group and control. Experiments were analysed using student t-test.

the presence of inhibitors) undergo apoptosis which stunted the growth of the tail at the region of damage, while the healthy cells adjacent to them grow normally, producing the bent phenotype.

For both the curved-body and bent floor-plate phenotypes, none of the control group embryos had developed these phenotypes (Figure 3C and 3D).

The manifestation of these phenotype was of the same pattern across the range of inhibitors' concentrations. In both Mirin and Ku55933 treated embryos, as the concentration increases, the rate of embryos developing these phenotypes increased with 50% being the highest for the curved body phenotype and 33% for the bent floor plate phenotype. However, embryos treated with Mirin at 12 µM did not develop any of these phenotypes and the embryos treated with Ku60019 had decreasing rate of the phenotype manifestation as the concentration increases. The data indicates the severity of these two phenotypes is independent of the concentration of inhibitors used.

Our findings indicate that embryos can survive and hatch phenotypically normal when treated with inhibitors alone. However, with the occurrence of DNA damage inflicted by the UVC radiation, majority of the inhibitor treated embryos that are alive pass the hatching stage, develop into phenotypically deformed embryos. This data is statistically significant when compared to the control group.

CONCLUSION

We found that specific protein inhibitors such as Mirin, Ku60019 and Ku55933 increased the possibilities of deformities formation particularly in enlarged yolk sac, curved-body and bent floor-plate phenotypes among the UVC treated embryos. In the presence of protein inhibitors however, we observed that embryos without UVC exposure did not show any abnormal phenotypes. This suggests that RAD50 and ATM do not involve directly in zebrafish embryogenesis but may be crucial in safeguarding the whole embryogenesis process as a whole.

ACKNOWLEDGEMENTS

This work was supported by the Ministry of Science, Technology and Innovation (MOSTI) of Malaysia (Grant no.: ScienceFund 02-01-02-SF1279), the Ministry of Higher Education (MOHE) of Malaysia (Grant no.: FRGS/2/2014/ST04/UKM/03/1) and Pusat Penyelidikan Tasik Chini (PPTC), UKM (LIV-2015-05).

REFERENCES

Bender, C. F., Sikes, M. L., Sullivan, R., Huye, L. E., Le Beau, M. M., Roth, D. B., … & Petrini, J. H. J. (2002). Cancer predisposition and hematopoietic failure in Rad50S/S mice. *Genes and Development, 16*(17), 2237-2251.

Bladen, C. L., Lam, W. K., Dynan, W. S., & Kozlowski, D. J. (2005). DNA damage response and Ku80 function in the vertebrate embryo. *Nucleic Acids Research, 33*(9), 3002-3010.

Chwastek, J., Jantas, D., & Lasoń, W. (2017). The ATM kinase inhibitor KU-55933 provides neuroprotection against hydrogen peroxide-induced cell damage via a γH2AX/p-p53/caspase-3-independent mechanism: Inhibition of calpain and cathepsin D. *International Journal of Biochemistry and Cell Biology, 87*, 38-53.

Clemente, Z., Castro, V. L. S. S., Moura, M. A. M., Jonsson, C. M., & Fraceto, L. F. (2014). Toxicity assessment of TiO_2 nanoparticles in zebrafish embryos under different exposure conditions. *Aquatic Toxicology, 147*, 129-139.

Dale, R. M., Sisson, B. E., & Topczewski, J. (2009). The emerging role of Wnt/PCP signalling in organ formation. *Zebrafish, 6*(1), 9-14.

Driever, W., Schier, A. F., Neuhauss, S. C. F., Malicki, J., Stemple, D. L., Stainier, D. Y. R., … & Boggs, C. (1996). A genetic screen for mutations affecting embryogenesis in zebrafish. *Development, 123*, 37-46.

Dubois, E., & Gerber, S. (2016). UV-dependent phosphorylation of COP9/signalosome in UV-induced apoptosis. *Oncology Reports, 35*(5), 3101-3105.

Dupré, A., Boyer-Chatenet, L., Sattler, R. M., Modi, A. P., Lee, J. H., Nicolette, M. L., & Gautier, J. (2008). A forward chemical genetic screen reveals an inhibitor of the Mre11–Rad50–Nbs1 complex. *Nature Chemical Biology, 4*(2), 119-125.

Fazry, S., Noordin, M., Sanusi, S., Noor, M., Aizat, W., Lazim, A., … & Law, D. (2018). Cytotoxicity and toxicity evaluation of xanthone crude extract on hypoxic human hepatocellular carcinoma and zebrafish (*Danio rerio*) embryos. *Toxics*, *6*(4), 60-70.

Gatei, M., Jakob, B., & Chen, P. (2011). ATM-dependent phosphorylation of Rad50 regulates DNA repair and cell cycle control. *Journal of Biological Chemistry, 286*(36), 31542-31556.

Golling, G., Amsterdam, A., Sun, Z., Antonelli, M., Maldonado, E., Chen, W., … & Hopkins, N. (2002). Insertional mutagenesis in zebrafish rapidly identifies genes essential for early vertebrate development. *Nature Genetics, 31*(2), 135-140.

Kaur, S., Kaushik, S., & Kedarisetti, K. C. (2018). Lasers in glaucoma. In H. V. Nema & N. Nema (Eds.), *Gems of Ophthalmology: Glaucoma* (pp. 334-354). New Delhi, India: Jaypee Brothers Medical Publisher.

Lavin, M. F. (2007). ATM and the Mre11 complex combine to recognize and signal DNA double-strand breaks. *Oncogene, 26*(56), 7749-7758.

Ledo, E. (1993). Photodermatosis part I: photobiology, photoimmunology, and idiopathic photodermatoses. *International Journal of Dermatology*, *32*(6), 387-396.

Luo, G. (1999). Disruption of mRad50 causes embryonic stem cell lethality, abnormal embryonic development, and sensitivity to ionizing radiation. *Proceeding of the National Academy of Sciences, 96*(13), 7376-7381.

Meyers, J. R. (2018). Zebrafish: development of a vertebrate model organism. *Current Protocols Essential Laboratory Techniques, 16*(1), 1-26.

Mullenders, L. H. (2018). Solar UV damage to cellular DNA: from mechanisms to biological effects. *Photochemical and Photobiological Sciences, 17*(12), 1842-1852.

Nair, S. G., & Loppnow, G. R. (2019). Comparison of K-Ras and N-Ras mutagenic hot spots for UVC damage. *ACS Omega*, *4*(2), 3469-3475.

Negi, S. S., & Brown, P. (2015). rRNA synthesis inhibitor, CX-5461, activates ATM/ATR pathway in acute lymphoblastic leukemia, arrests cells in G2 phase and induces apoptosis. *Oncotarget, 6*(20), 18094-18104.

Segner, H. (2009). Zebrafish (*Danio rerio*) as a model organism for investigating endocrine disruption. *Comparative Biochemistry and Physiology Part C: Toxicology and Pharmacology, 149*(2), 187-195.

Sugiyama, T., & Chen, Y. (2019). Biochemical reconstitution of UV-induced mutational processes. *Nucleic Acids Research, 2019*, 1-14.

Waltes, R., Kalb, R., Gatei, M., Kijas, A. W., Stumm, M., Sobeck, A., … & Dörk, T. (2009). Human RAD50 deficiency in a Nijmegen breakage syndrome-like disorder. *The American Journal of Human Genetic, 84*(5), 605-616.

Williams, B. R., Mirzoeva, O. K., Morgan, W. F., Lin, J., Dunnick, W., & Petrini, J. H. J. (2002). A murine model of Nijmegen breakage syndrome. *Current Biology, 12*(8), 648-653.

Williams, T. M., Nyati, S., Ross, B. D., & Rehemtulla, A. (2013). Molecular imaging of the ATM kinase activity. *International Journal of Radiation Oncology Biology Physics, 86*(5), 969-977.

Zhu, J., Petersen, S., Tessarollo, L., & Nussenzweig, A. (2001). Targeted disruption of the Nijmegen breakage syndrome gene NBS1 leads to early embryonic lethality in mice. *Current Biology, 11*(2), 105-109.

Pertanika J. Sci. & Technol. 27 (4): 1625 - 1647 (2019)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Identifying Analogues Of 2-Deoxyglucose, Alpha-D-Glucose and Beta-D-Glucose-6-Phosphate as Potential Inhibitors of Human Hexokinase II for the Development of Anti-Dengue Therapeutics

Suriyea Tanbin[1], Nurhainis Ogu Salim[2] and Fazia Adyani Ahmad Fuad[1]*

[1]Department of Biotechnology Engineering, Faculty of Engineering, International Islamic University Malaysia, 50728 Kuala Lumpur, Malaysia
[2]Parasitology Unit, Institute for Medical Research, Setia Alam, 40170 Shah Alam, Selangor, Malaysia

ABSTRACT

The human hexokinase isoform II (HKII) is one of the important enzymes for dengue virus (DENV) replication and thus has been suggested as a potential therapeutic target for DENV drug development. In this work, compounds were identified using Ultrafast Shape Recognition with CREDO Atom Types (USRCAT) by utilizing both HKII's substrate and product; alpha-D-glucose (GLC) and beta-D-glucose-6-phosphate (BG6), as well as a known HKII's inhibitor, 2-deoxyglucose (2DG), as the query molecules. The analogues of the three query molecules were subsequently docked against the HKII's crystal structure (PDB ID: 2NZT) by using Auto Dock 4 program on Chain B, where the active sites and strong bonds were located. Among the top-ranked compounds, Compound 4 (ZINC26898487), which was the most similar to 2DG, showed the best binding energy (-7.63 kcal/mol) and contained two H bonds. Compound 9 (ZINC16930948), an analogue of GLC emerged as the best inhibitor candidate because it had six H bonds. Similarly, among the molecules similar to BG6, Compound 14 (ZINC4403351) had been suggested as a potential inhibitor because it contained four strong H bonds. All compounds were predicted to be non-toxic, based on Toxicity Estimation Software Tool (TEST) analysis. By providing these valuable findings, this study has paved the way for the discovery of compounds that should be further tested for the development of anti-dengue drugs.

Keywords: 2-deoxyglucose, alpha-D-glucose, beta-D-glucose-6-phosphate, Human Hexokinase II (HK2), ligand-based screening, structure-based screening, toxicity test

ARTICLE INFO

Article history:
Received: 18 May 2019
Accepted: 19 July 2019
Published: 21 October 2019

E-mail addresses:
tanbin.khu.bd@gmail.com (Suriyea Tanbin)
Nurhainis@gmail.com (Nurhainis Ogu Salim)
fazia_adyani@iium.edu.my (Fazia Adyani Ahmad Fuad)
* Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

INTRODUCTION

The world is currently experiencing a steady increase in the number of dengue-related cases, and the prevalence of this disease remains high (World Health Organization, 2018). Approximately 390 million people are infected with dengue virus every year, resulting in around 25,000 annual deaths (World Mosquito Program, 2018). At present, this disease is registered in more than 128 countries worldwide, with the greatest burden imposed on the South-East Asian, Eastern Mediterranean and South American countries (World Health Organization & World Mosquito Program, 2018). In Western Pacific region the greatest dengue incidence was noted in Malaysia, Vietnam and Philippines, with 41,443 cases (65 deaths), 52,482 cases (3 deaths) and 59,139 cases (237 deaths), respectively as of week 16, 2019 (World Health Organization, 2019).

Mosquito-borne dengue virus is responsible for the intense feverish dengue disease, whereby female *Aedes aegypti* or *Aedes albopictus* mosquito acts as a vector for virus transmission to humans via a single bite (Wichapong et al., 2013). Dengue virus is a single-stranded RNA virus belongs to Flaviviridae family, which exhibits four serotype types; DENV-1, DENV-2, DENV-3, and DENV-4 (Diamond & Pierson, 2015). The severity of the dengue disease is determined by sequential attacks by different serotypes.

Previously, it has been established that DENV-1 is more virulent compared to DENV-2 based on the clinical manifestation, hematological parameters and genotypic variation of limited dengue cases (Fried et al., 2010; Yung et al., 2015) . However, at present, DENV-2 is considered the most virulent serotype based on the clinical data and laboratory test reports (Idrees et al., 2012; Vinodkumar et al., 2013; Mudin, 2015; Yung et al., 2015; Wijewickrama et al., 2018). According to the published data, the percentage of DENV-2 serotypes present in patients' blood ranges from 96.02% (Idrees et al., 2012) to 92.4% (Wijewickrama et al., 2018), while 29.2% DENV-2 serotype was also reported from blood samples of 72 patients (Vinodkumar et al., 2013). However, although serotypes share approximately 65% of the genomic character, they remain difficult to distinguish (Yung et al., 2015; Mirza et al., 2016).

Unfortunately, at present, there is no effective vaccines or specific drugs that can be integrated into the global dengue prevention and control strategy (World Health Organization, 2018). However, as a part of their recent clinical study, Godói and research team (2008) produced the first dengue vaccine, Dengvaxia (CYD-TDV), which was registered in Mexico, by Sanofi Pasteur. However, at present, the vaccine has drawbacks since it cannot provide long-term protection against all four serotypes, and it is only effective for children aged 2−16 years (Godoi et al., 2008). It has also been established that sequential infections with different serotypes increase the chances of developing severe disease, which limits the use of vaccination as an early step in dengue prevention. As noted above, no well-developed antiviral drugs are presently available, which could cure this disease.

Thus, the treatment mostly relies on rehydration therapy. Typically, rehydration therapy consists of consuming high protein, low fat, non-oily, and non-spicy diet and drinking lot of water, as this has been found to increase appetite and improve nutritional status, while also elevating the platelet count, thus improving the patient's immunity (Mishra et al., 2017). However, dengue patients can suffer from wide range of health complications, such as compromised food digestion, low nutritional status, and immunity reduction, all of which can benefit from rehydration therapy. While these measures can alleviate dengue symptoms, there is an urgent need to discover new potentially effective drugs that may serve as the first line of treatment for dengue fever. As viruses are inanimate entities and do not possess metabolic machinery, they rely on their host for survival and replication (Wick et al., 1956; Schreyer & Blundell, 2012; Fontaine et al., 2015). During DENV infection, notable glucose consumption was observed by Fontaine and colleagues, who also noted that absorption of exogenous glucose had a significant impact on DENV replication. This consumption process mainly occurred in the glycolytic pathway (Fontaine et al., 2015). The authors further observed that, the expression of hexokinase isoform II (HKII) was significantly higher in the virus-affected cell, compared to the mock-infected cell. It is also important to highlight that hexokinase is the first rate-limiting enzyme of glycolytic pathway and HKII is found to play a key role of aerobic glycolysis (Gershon et al., 2013). In this pathway, with hexokinase assistance, glucose is converted to glucose-6-phosphate (G6P). G6P is produced by the phosphorylation of glucose, where hexokinase enzyme catalyzes this reaction by consumption of adenosine triphosphate (ATP). The reason behind glucose phosphorylation is that phosphorylated glucose cannot easily cross the cell membrane since negative phosphate group is attached with glucose, whereby phosphorylation traps glucose in the cell (Lunt & Heiden, 2011; Pelliccia et al., 2017). As mentioned above, viruses are lacking of metabolic function, thus they take the biosynthetic element and energy from host cell through the modulation of human's metabolic pathway. Several experiments have been conducted to observe the alteration of host cell metabolic pathway during viral infection, for example on vaccinia virus (VACV), human cytomegalovirus (HCMV) and hepatitis C virus (Munger et al., 2006; Roe et al., 2011; Fontaine et al., 2014). All these viruses caused extensive reprogramming of central carbon metabolism during infection, and similar to tumor and cancer cell, HCMV-infected cell absorbed glucose through glucose transporter 4 (GLUT4) and partially breakdown the glucose carbon through TCA cycle. Subsequently, fatty acid would be synthesized, which is crucial for HCMV replication (Yu et al., 2011). As a substitute, VACV virus may need another carbon source glutamine for the production of virion through maintenance of TCA cycle (Fontaine et al., 2014). During DENV infection, the infected cells increase the up-regulation of glucose through GLUT1 and GLUT4, as well as HKII, where later on glucose is diverted to cytoplasm through glycolytic pathway and TCA cycle, that generates a bulk of ATP and fatty acid biosynthesis required for viral protein synthesis as well as proliferation. Glucose consumption of DENV is thus measured by highly significant HKII expression level. Pelliccia and colleagues (2017) had recently shown that DENV replication and glycolysis were positively correlated, suggesting that the inhibitory treatment of glucose competitor, 2-deoxyglucose (2DG), could significantly hamper DENV replication. Wick and colleagues (1956) established that 2DG was an analogue of glucose, in which two hydroxyl groups were replaced by hydrogen, which inhibited the production of glucose-6-phosphate and eventually hampered the glycolytic pathway (Godoi et al., 2008). Later, 2DG has been suggested by several work as a glycolytic inhibitor in cancer, tumor, as well as dengue virus replication (Ganapathy-Kanniappan & Geschwind, 2013; Kaushik et al., 2015; Fontaine et al., 2015; Savic et al., 2016). However, 2DG has shown toxic effects when introduced to patients (Zhang et al., 2014). Hence, it is important to identify new compounds similar to 2DG that would not produce such toxicity effects, as these could be valuable candidates for anti-dengue drug development. This was the goal of the present investigation, as a part of which the inhibition activity of 2DG was considered. Apart from 2-deoxyglucose (2DG), both alpha-D-glucose (GLC) and beta-D-glucose-6-phosphate (BG6) serve as interesting reference molecules for rational drug design where GLC, as the substrate for HKII-catalyzed reaction, is an anomer of glucose-6-phosphate.

In this work, we report the identification of small molecules similar to 2DG, GLC, and BG6, which have the potentials to be developed into safe and effective anti-DENV drugs. A ligand-based drug design experiment was conducted by screening similar compounds that were analogues of the substrates and product of human HKII, namely alpha-D-glucose (GLC), and beta-D-glucose-6-phosphate (BG6) as well as 2DG, which was the known inhibitor of HKII. Subsequently, a series of docking and scoring experiments was conducted, following a structure-based drug design screening based on the crystal structure of human HKII, before these compounds were being tested in virtual toxicity test.

MATERIALS AND METHODS

Ligand-Based Screening

As previously noted, 2DG, GLC and BG6 were used as query molecules for the initial ligand-based pharmacophore screening (Figure 1). The 3D structures of 2DG, GLC and BG6 (with PubChem ID CHEBI: 10822 ID CHEBI: 17925 and ID CHEBI: 17719, as well as molecular formula $C_6H_{12}O_5$ $C_6H_{12}O_6$ and $C_6H_{13}O_9P$), were downloaded in structure data file (SDF) format from PubChem (https://pubchem.ncbi.nlm.nih.gov/) to screen for similar compounds by using USRCAT (Ultrafast Shape Recognition with CREDO Atom Types) available at http://usr.marseille.inserm.fr/. We retrieved a total of 300 compounds and their ZINC IDs, where each of the query molecule (2DG, GLC and BG6) had 100 similar hits from the USRCAT program, whereas the 2D structure, hydrogen bond donor, hydrogen bond acceptor, molecular weight, and logP value of every hit can be found in the ZINC (http:/zinc.docking.org/) database. The predicted similarity of the molecules was based on the scoring range 0 < score < 1, where a compound with higher score indicated greater similarity towards its respective query molecule. Based on the similarity score, the top 40 compounds for each of the query molecules (2DG, GLC, and BG6) were chosen, thus a total 120 compounds were selected for further docking studies.

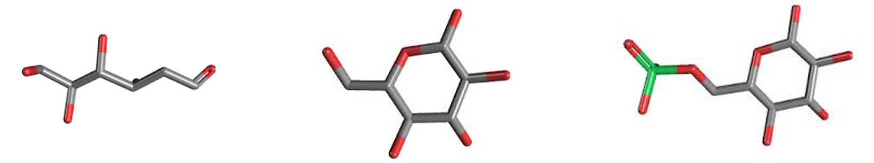


Figure 1. The 3-dimensional structure of 2-deoxyglucose (2DG), alpha-D-glucose (GLC), and beta-D-glucose-6-phosphate (BG6) molecule used in the ligand-based virtual screening analysis (PubChem ID CHEBI: 17925, ID CHEBI: 10822 and ID CHEBI: 17719). The grey, red, and green color indicate carbon, oxygen, and phosphate atoms, respectively

Structure-Based Screening

For the present study, the crystal structure of human HKII was retrieved from the Protein Data Bank (PDB ID: 2NZT) website https://www.rcsb.org/. The protein structure was prepared for molecular docking by protein optimization and energy minimization, which was performed by removing ligands from the protein structure using SPDBV-Swiss-PdbViewer software. Molecular docking was conducted by using two different software tools, namely Auto Dock 4 and Cygwin program, whereby Auto Dock 4.2 was used to dock each compound into the target protein and Cygwin program assisted in performing the docking procedure by using genetic and Lamacrkian algorithm (Morris, et al., 2010). We prepared a grid box that encompassed the binding site of the protein, where ligands were docked sequentially. The grid box covered the predicted catalytic protein residues, where the box size was (40 × 40 × 40 Å) and the box center coordinates (x = 13.062; y = 13.881; z = 7.683), on the C terminal of the protein's chain B. Cygwin was used for the docking algorithm, in which grid log and docking log files were generated. The binding energy (kcal/mol), derived from the docking log file, indicated the strength of the ligand-receptor binding, which was subsequently analyzed by using PyMol software.

Toxicity Test

Toxicity estimation software (TEST) version 4.2.1 was used to determine the toxicity of the top-ranked six compounds, following docking analysis. The protein data file (pdb file) format was converted into a molecular data file (MDL mol) through Open Bable software. We used Quantitative Structure Activity Relationship (QSARs) methodologies to predict toxicity measures from physical characteristics, based on the compounds' structure. Selected compounds were tested using different test end-points, such as lethal dose LC50, bioconcentration factor (BCF), developmental toxicity, and mutagenicity.

RESULTS AND DISCUSSION

Ligand-Based Screening Analysis

Ligand-based virtual screening has been proven as an effective tool for discovering novel chemical platforms in compound libraries (Vyas et al., 2008; Gao et al., 2010; Lavecchia & Giovanni., 2013). In the present study, a search was carried out by using both HKII's substrate and product (GLC and BG6, respectively), as well as known inhibitors of HKII (2DG) as the query molecules, to search for similar candidates by using USRCAT program. USRCAT is a user-friendly, cost-effective and time-effective web based program which is performed for large scale ligand-base screening. The USRCAT was developed as a part of USR algorithm to assist users in their search for desired compounds similar to the query molecules, that is able to screen more than 50 million of 3D conformers per second (Schreyer & Blundell, 2012). The latest version of USRCAT is capable of differentiating not only the 3D shapes, but also the distribution of atom types pertinent for query molecule recognition (aromatic, hydrogen bond donor, hydrogen bond acceptor, and hydrophobic atoms). As a part of the current investigation, 23.1 million molecules were screened and a total of 300 most similar compounds were obtained from USRCAT for all of the reference molecules (2DG, GLC, and BG6). Thus, specifically, each query has 100 similar compounds aligned into the hits.csv file along with similarity scores, physicochemical properties, and the vendor's information. In this work, similarity was evaluated based on the scoring function range from 0-1, where zero and one indicates minimum and maximum similarity, respectively. As mentioned above, the USRCAT score zero indicates the least similar compounds, while those closer to one reflects a much closer resemblance towards the reference molecule. Based on these findings, a total 120 compounds out of 300 similar hits for each query molecule were chosen and their physiological properties were obtained from the ZINC database. The ZINC database contains 727,842 purchasable compounds from various suppliers (ChemBridge, ChemDiv, Ryan, Asinex, Sigma-Aldrich, Maybridge, Specs, Comgenex and Otava), 494,915 of which are Lipinski subjacents. Based on the similarity scores, the six top-ranked compounds that resembled each of the query molecule are shown in Table 1-3, as their ranking scores ranged from 0.76 to 0.79, 0.85 to 0.88 and 0.84 to 0.86 for 2DG, GLC and BG6 respectively. All six compounds, which are analogues of 2DG, GLC, and BG6 which have been selected for the subsequent studies obeyed the Lipinski's Rule of Five in terms of their physiological parameters; hydrogen bond donors > 5, hydrogen bond acceptor > 10, molecular weight under > 500g/mol, partition coefficient logP > 5 with no violation (Lipinski et al., 1997). The mentioned parameters related to the Lipinski's rule of five are presented in Tables 1 to 3. Compounds 4 to 21, which are similar conformers of 2DG possess MlogP values ranging from -1.73 to -2.34, and drug-likeness from -0.92 to -1.21 (Table 1). Moreover, the top six similar hits of GLC, compounds 8 to 26 have MlogP values from -3.12 to -3.04 and drug-likeness from -1.21 to -0.15 (Table 2). Drug-likeness is the preliminary concept of drug design and its value is usually estimated from the compound's molecular structure that is standardized between -0.4 and 5.6, whereby it indicates whether the compound has drug properties or not. As our findings were optimized with respect to the drug-likeness parameter and the adherence to

Table 1
Physicochemical properties and similarity scores for 2DG and its analogues, as obtained from USRCAT

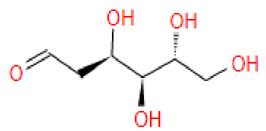
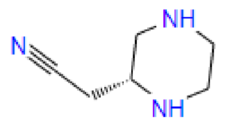
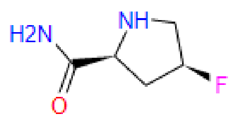
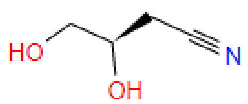
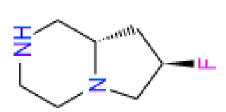
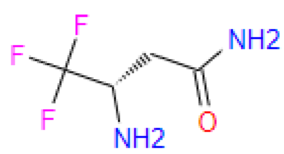
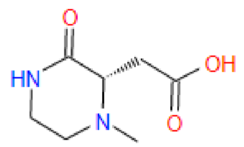
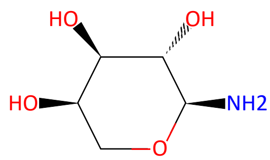
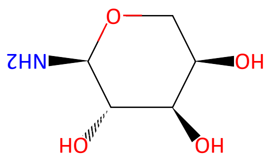
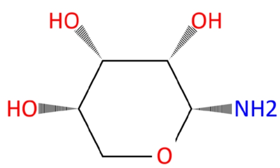
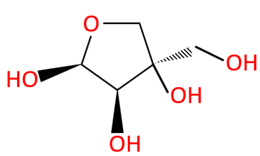
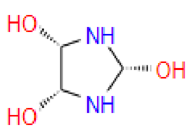
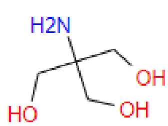
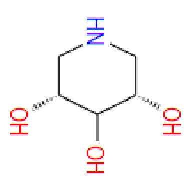
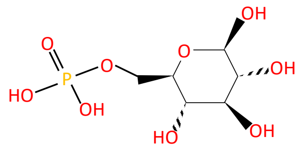
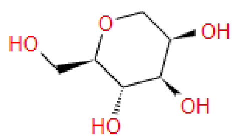
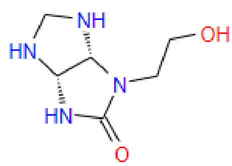
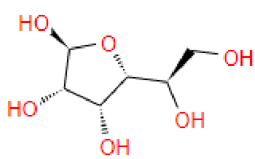
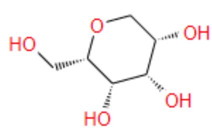
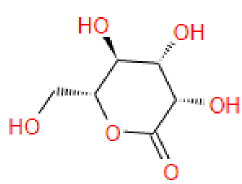
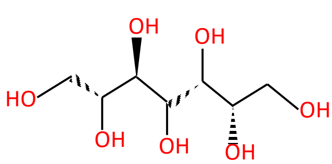
Compound number	2D structure	Similarity score	Number of HBD/HBA	Molecular weight (g/mol)	MlogP	Drug-likeness
2DG		N/A	4/5	164.157	-1.20	-0.50
4		0.79	2/3	125.175	-1.73	-0.92
6		0.78	3/3	132.138	-1.66	-0.67
8		0.78	2/3	101.105	-1.44	-1.39
18		0.76	2/2	145.201	-1.64	-1.16
19		0.76	4/3	156.107	-1.34	-1.19
21		0.76	2/5	172.184	-2.34	-1.21

Table 2
Physicochemical properties and similarity scores for GLC and its analogues, as obtained from USRCAT

Compound number	2D structure	Similarity score	Number of HBD/HBA	Molecular weight (g/mol)	MlogP	Drug-likeness
GLC		N/A	5/6	180.156	-2.11	-0.20
8		0.88	5/5	149.146	-3.12	-1.21
9		0.88	5/5	149.146	-3.12	-1.21
12		0.87	4/5	150.13	-2.38	-0.15
18		0.86	5/4	134.155	-2.77	-1.45
25		0.85	6/4	122.144	-2.20	-0.70
26		0.85	5/5	120.108	-3.04	-1.02

the Lipinski's Rule of Five, most of the compounds fulfilled all the criteria. The drug-like properties of BG6 corresponded to the MlogP values between -2.57 and -4.26, and drug-likeness between -0.92 and 0.12 (Table 3). It is noteworthy to note that, although the same scores were obtained from USRCAT program, these compounds are structurally different and contain different atom types. The rule of five provides insights for the analysis of the

Table 3
Physicochemical properties and similarity scores for BG6 and its analogues, as obtained from USRCAT

Compound number	2D structure	Similarity score	Number of HBD/HBA	Molecular weight (g/mol)	MlogP	Drug-likeness
BG6		N/A	2/2	152.132	-2.07	-0.42
2		0.86	4/5	146.157	-2.57	-0.92
3		0.86	4/6	172.188	-2.68	-1.10
12		0.85	5/6	180.156	-3.05	-0.63
14		0.85	4/5	164.157	-2.57	-0.92
21		0.84	4/6	178.14	-3.45	-0.18
23		0.84	7/7	212.198	-4.26	0.12

molecular properties and structural features of every compound, which is important for drug pharmacokinetics in human body, including their absorption, distribution, metabolism and excretion (ADME) (Lipinski et al., 1997).

Computational methods are effectively used in pharmaceutical research to improve drug discovery process. Ligand-based screening is one of the crucial parts of *in silico* drug design and several novel compounds have been invented through this process. Hence, considering the importance of computational drug design method, it was adopted in the present study to discover the potential HKII inhibitors. HKII is an important drug target for designing novel compounds for dengue treatment. Numerous successful ligand-based screening investigations have been performed, aiming to identify potential inhibitors against dengue drug targets. In extant studies, dengue virus encoded non-structural protein 3(NS3) was used a as drug target, and two known inhibitors, namely suramin and ivermectin have been successfully identified through virtual screening and *in vitro* analysis, where researchers used LOPAC compound library for inhibitor selection as well as vero-B and vero-E6 cell for *in vitro* assay (Mastrangelo et al., 2012; Basavannacharya & Vasudevan, 2014). It is noteworthy to note that most of the research conducted to search for anti-dengue drugs have been conducted against viral replication proteins or viral envelop proteins, but our research focus is on human protein as a drug target. Another fruitful ligand-based work has also been done using Ultra-Fast Shape Recognition with Atom Type (UFSRAT) to identify similar hits of S-adenosyl-L-methionine (SAH) and ribavirin triphosphate (RTP), in which those were used against drug target NS5 methyltransferase (MTase). In their work, Lim and colleagues (2011) had identified 500 compounds that exhibited the highest similarity to SAH and RTP, which were subsequently used in docking analysis (Lim et al., 2011). Nevertheless, UFSRAT is also a powerful web program, but it shows complexity in selecting the best compounds without the 3D structure of the drug target. In our study, only the 3D structure of the reference molecule has been used for screening and a total 120 compounds were chosen for further docking analysis. Ligand-based methods not only yield effective results pertinent to dengue research, but are also beneficial in cancer or tumor studies. For instance, wild-type p^{53} protein is responsible for inducing cancer, but its function is inhibited when in complex with human murine double minute 2 (MDM2), which is why p^{53}-MDM2 is an important drug target (Patil et al., 2014). Basically, ligand-based screening is a very convenient process to discover drug-like molecules within a short period of time. There are numerous web-based programs available to implement this process, whereas in this work, an on line program, USRCAT has been utilized, which is incorporated with ZINC database. Lipinski's Rule of Five functions as a filter to choose drug-like compounds based on their molecular property, including absorption, distribution, metabolism and excretion. Hence, considering the importance of computational drug design method, it was adopted in the present study to discover the potential HKII inhibitors. Findings yielded have shown selected compounds, based on the similarity score and drug-like criteria.

Structure-Based Screening Analysis

Structure-based virtual screening is an effective process that has become instrumental in the fast and cost-effective lead discovery and optimization. The goal of this process is to understand the disease at a molecular level and apply the knowledge of three-dimensional structure of the biological target (Kroemer, 2007; Lionta et al., 2014). Structure-based screening refers to the protein-ligand binding interaction, where the protein-ligand binding mode can be predicted through docking studies, which consist of two-step calculation involving the free binding energy of protein-ligand complex, and the conformational space exploration of ligand and target molecules (Reddy et al., 2007; Lavecchia & Giovanni, 2013). Prior to the emergence of the docking method, the structure was typically validated by redocking experiments, whereby the ligand was extracted from protein−ligand complexes to minimize the conformation and was subsequently docked back into the protein to validate the docking location (Cosconati et al., 2010).

In the present study, a successful docking site was determined by Root Mean Square Deviation (RMSD) value from known conformation, whereby a good docking site is generally indicated by an RMSD ≤ 2 Å (Nissink et al., 2002; Vyas et al., 2008).According to the RMSD value, the most accurate sites were selected by applying a scoring and ranking function. In this research, the 3D structure of human HKII (PDB ID: 2NZT) was retrieved from Protein Data Bank, in complex with alpha-D-glucose (GLC) and beta-D-glucose (BG6), which were removed prior to performing the docking experiment. Protein energy minimization is essential to set a protein structure in energetically-favorable state (reduction in the relaxing bond lengths, angles, and non-bond interactions), which provides more accurate docking results. Protein optimization and energy minimization were performed in the present study to avoid the complexity associated with protein−ligand docking. After protein preparation, the docking procedure was applied for 120 compounds that are identified from ligand-based screening. Each compound had been successfully docked into the catalytic site, which was validated through re-docking process as well as RMSD value. As mentioned, HKII structure in complex with the known ligands BG6 and GLC was utilized to validate the docking process, by extracting these two ligands, re-docking the ligands into HKII, prior to docking the virtually-screened compounds into the catalytic site of HKII. After completion, successful docking had been observed, where in the catalytic residues, the dock poses of the compounds were similar to the reference molecule and the RMSD value was reasonable (Figure 2, 3 and 4).

In molecular docking, the grid box size and the coordinates of its center have to be validated, in order to ensure that the ligands bind to the binding pocket in the correct conformation. The grid box (Box size: 40 × 40 × 40 Å) and the box center coordinates (x = 13.062; y = 13.881; z = 7.683) were designed, whereby all binding modes were created within this dimension for the most favorable bindings. The Human HKII (PDB ID: 2NZT) is a homodimer protein consisting of Chain A and B with catalytic C-terminal and N-terminal domains. In this work, we used C-terminal domain of Chain B for docking because several extant studies had revealed that the C-terminal (residues 476−917) contained more capable catalytic sites (Printz et al., 1997; Bianchi et al., 1999). In the present study, Thr620, Lys621, Glu629, Asn656, Asp657, Phe623, Ser893, Asn683, Gln739, and Glu742 (Figure 2, 3 and 4) catalytic residues were predicted as the most favorable sites for docking the compounds. The docking process involved 120 compounds, with the selected residue of catalytic sites for each of the query molecule (2DG, GLC, and BG6). All the lead molecules successfully docked into the catalytic site of human HKII protein. In this study, the six top-ranked compounds (4, 18, 6, 8, 19 and 21), which are analogues of 2DG had the binding energy ranging from -7.63 kcal/mol to -4.98 kcal/mol, whereas the binding energy (kcal/mol) of 2DG itself is -7.40 kcal/mol (Table 4). Among all conformers of 2DG, compound 4 (ZINC2689487) poses good binding energy (-7.63 kcal/mol) with two H bonds attached with residues Ser893 and Gln739, that closely resembles the query molecule 2DG (-7.40 kcal/mol), which poses six H-bonds; four attached with Ser893 ,one with Asn656 and another one is Thr620 (Figure 2B and 2A, respectively). It can be suggested that compound 4 (ZINC2689487) was the best findings among similar analogues of 2DG and might be a potential inhibitor of HKII. Furthermore, compounds 8, 9, 25, 26, 18 and 12, which are the analogues of GLC exhibited binding energy ranged from -8.87 kcal/mol to -7.16 kcal/mol, whereas GLC itself poses a binding energy of -6.40 kcal/mol (Table 5). Compound 9 (ZINC16930948) had been chosen as the best analogue based on highest binding affinity value (-8.50 kcal/mol) with six H bonds, four of which are formed with Thr620, Asn683, Asp657 and Glu742, and the remaining two with Asn656 (Figure 3B), when compared to the minimum binding energy of GLC in HKII (-6.40 kcal/mol) (Table 5), and it forms up to two H bonds with Glu742 and Asn656 (Figure 3A). The best compound among GLC analogues was Compound 9 (ZINC16930948), which might be a potential inhibitor of HKII.

On the other hand, similar hits of BG6 (compounds 23,3,2,21,12 and 14) exhibited binding energy ranged from -9.16 kcal/mol to -7.09 kcal/mol, as shown in Table 6. It can be noted that Compound 14(ZINC4403351) exhibited relatively good binding energy (-7.09 kcal/mol) with four H bonds, three of which were formed with Glu629 and one with Asn656 (Figure 4B), compared to binding energy of BG6 (-5.02kcal/mol) with one H bond with Phe623 (Figure 4A). Compound 14 (ZINC4403351) thus had been suggested for further inhibition analysis, due to having good number of H bonds, as well relatively good binding affinity, compared to BG6 itself and all other analogues of BG6.

Docking sites were considered successful based on the value of three parameters. First, re-docking experiments were conducted, by setting the RMSD = 2 Å for all docked sites, as this had been considered as a threshold value. Secondly, predicted catalytic residues were present in the protein structure, before estimating the minimum binding energy and including the number of H bonds with the selected residues into the catalytic sites of the protein. The catalytic sites of human HKII were determined based on the selected residues. In this case, the RMSD value of each dock site was 2 Å for every selected compound that was similar to each query molecules (2DG, GLC and BG6), and it was determined by known protein−ligand dock sites retrieved from the PDB file. The substrate and product (GLC and BG6) in complex with the protein were used as the primary lead for introducing these ligand derivatives. According to the findings yielded by other studies that have been conducted in the search for antiviral activity of potential inhibitors against NS2B/NS3 target protein *via* the docking method, the search was conducted using three query molecule such as mycophenolic acid (MPA), ribavirin and panduratin, (PubChem ID446541, ID37542 and ID6483648 respectively), which are the most promising known inhibitors (Parida & Yadav, 2014). However, in their study, Parida and Yadav (2014) utilized these three different query molecules, and the findings suggested that panduratin's derivatives were better inhibitors compared to mycophenolic acid and ribavirin derivatives, based on the docking score using PatchDock online server, which helped in the prediction of protein-ligand interaction and H-bond analysis results. The selected panduratin derivative's (compound 1) PatchDock score is 3698 and it contains four H bonds. In contrast, our results indicate that compound 4 (ZINC26898487), compound 9 (ZINC16930948) and compound 14 (ZINC4403351), the analogues of 2DG, GLC and BG6 respectively, are the most potential inhibitors, and should be subjected to further *in vitro* inhibition analysis based on their binding energy and H-bond analysis findings. Indeed, based on a comparison between the similarity hits and the query molecules, compound 4, the analogue of 2DG has good similarity score (0.79), binding energy (-7.63kcal/mol) and contains two H bonds (Tables 1 and 4) (Figure 2B). Meanwhile, Compound 9 and Compound 14 were chosen (Table 2, 3, 5 and 6) because of good similarity score (0.88) and (0.85), binding energy (-8.50 kcal/mol) with six H bonds (Figure 3B), and binding affinity (-7.09 kcal/mol) with four H bonds (Figure 4B), which are analogues of GLC and BG6, respectively. In this case, although other compounds also show relatively good similarity scores and binding affinities, but since these compounds lacked H bonds, they were not suggested for further inhibition studies.

Table 4
The top six molecules, which are analogues of 2DG that were ranked according to their binding energy. The binding energy of 2DG is also shown

Compound	ZINC ID	Binding energy (Kcal/mol)
2DG	2512351	-7.40
4	26898487	-7.63
18	39957300	-6.95
6	71257594	-5.97
8	2564305	-5.07
19	19795285	-5.07
21	1926703	-4.98

Table 5
The top six molecules, which are analogues of GLC that were ranked according to their binding energy. The binding energy of GLC is also shown

Compound	ZINC ID	Binding energy (Kcal/mol)
GLC	1456321	-6.40
8	35644927	-8.87
9	16930948	-8.50
25	896695	-8.33
26	15938179	-7.85
18	14658010	-7.78
12	14854290	-7.16

Table 6
The top six molecules, which are analogues of BG6 that were ranked according to their binding energy. The binding energy of BG6 is also shown

Compound	ZINC ID	Binding energy (Kcal/mol)
BG6	3875374	-5.02
23	17952732	-9.16
3	58951086	-7.69
2	5851505	-7.53
21	2539703	-7.28
12	34027427	-7.14
14	4403351	-7.16

In the previous work, *in silico* search for inhibitors against nonstructural protein 3 (NS3) by high- throughput virtual screening showed that, in descending order, ZINC95518765, ZINC44921800, ZINC71917414, ZINC39500661, and ZINC36681949 had the greatest binding capability with NS3, as well as the highest binding energies of -7.55, -7.36, -8.04, -8.41, and -9.18 kcal/mol, respectively. They also exhibited binding interaction with three catalytically important residues (His51, Asp75 and Ser135) of NS3 protein (Mirza et al., 2016). It should be noted that NS3 is a non-structural protein that assists in viral replication. In their computational study, Lim and colleagues (2011) searched for NS5 methyaltransferase (MTase) inhibitors by docking, reporting that compounds SPH1-103-799, SPH1-101-102 and 28SPH1-115-917 had the greatest binding affinity (-11.4 kcal/mol, -11.4 kcal/mol and -10.0 kcal/mol, respectively) (Lim et al., 2011). In the present study, Auto Dock Vina and EDULISS (EDinburg University Ligand Selection System) were used for structure-based and ligand-based screening, respectively. NS5 is a non-structural viral protein located on the dengue virus surface that assists in virus replication. Hence, it has been proposed as a drug target, whereby ribavirin triphosphate (RTP) and the S-adenosyl-L-methionine (SAM) act as query molecules. Four-stage computational HTS was used to find a small molecule against the envelope protein (E protein) of dengue virus (Wang et al., 2011). Next, 23 top-ranked compounds were selected based on the E protein binding site from three NIC libraries, which were tested for antiviral activity in biological assay. The compound PO2 was found to inhibit viral reproduction. Extensive research on the dengue virus has been conducted through the docking process (Wang et al., 2011; Yang et al., 2013; Behnam et al., 2014; Sing et al., 2015; Kouretova et al., 2017; Halim et al., 2017; Byrd et al., 2018). All the above mentioned literature reported inhibitors which were developed based on the viral protein itself. On contrary, our findings were based on human protein as a drug target. This structure-based process provides good assumption of protein-ligand binding affinity as well as assisting in choosing the best compound for further inhibition studies. Antiviral drug development relies on the appropriate compound selection through relevant *in silico* studies.

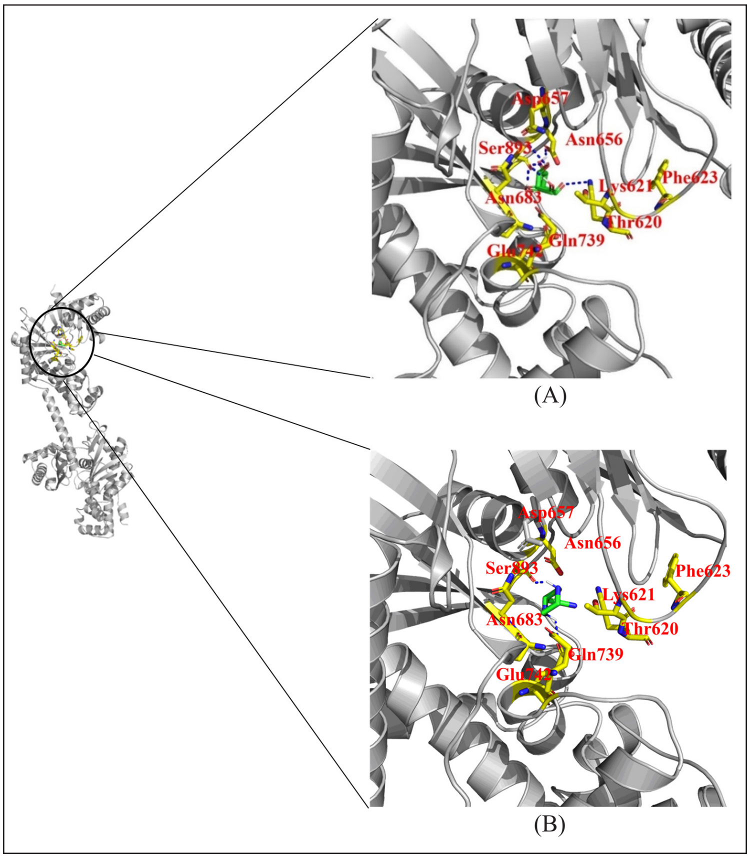


Figure 2. The interactions between 2DG and compound 4 with the residues involved in the catalytic site (A) HKII in complex with 2DG (green and red color indicates carbon and oxygen atom respectively of 2DG), which shows six H bonds; four with Ser893, one with Asn657 and another one formed with Thr620. (B) HKII in complex with compound 4 (green and blue color indicate carbon and nitrogen atom respectively of compound 4), which shows two H bonds formed with Ser893 and Gln739. Both interactions include the same catalytic residue (yellow and blue color indicates carbon and nitrogen atom of active sites residues) surrounding the ligand, while the blue lines indicate H bond.

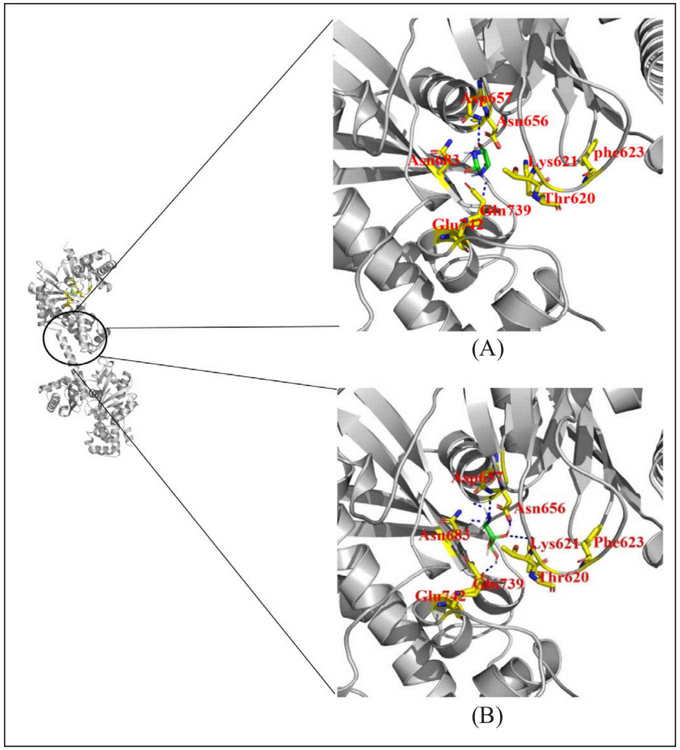


Figure 3. The interactions between GLC and compound 9 with the residues involved in the catalytic site (A) HKII in complex with GLC (green and blue color indicates carbon and nitrogen atom of GLC), which shows two H bonds formed with Glu742 and Asn656, respectively. (B) HKII in complex with compound 9 (green and red color indicate carbon and oxygen atom respectively of compound 9), which shows with six H bonds, each with Thr620, Asn683, Asp657 and Glu742 and another two H bonds formed with Asn656 (yellow and blue color indicates carbon and nitrogen atom of active sites residues) surrounding the ligand, while the blue lines indicate H bond.

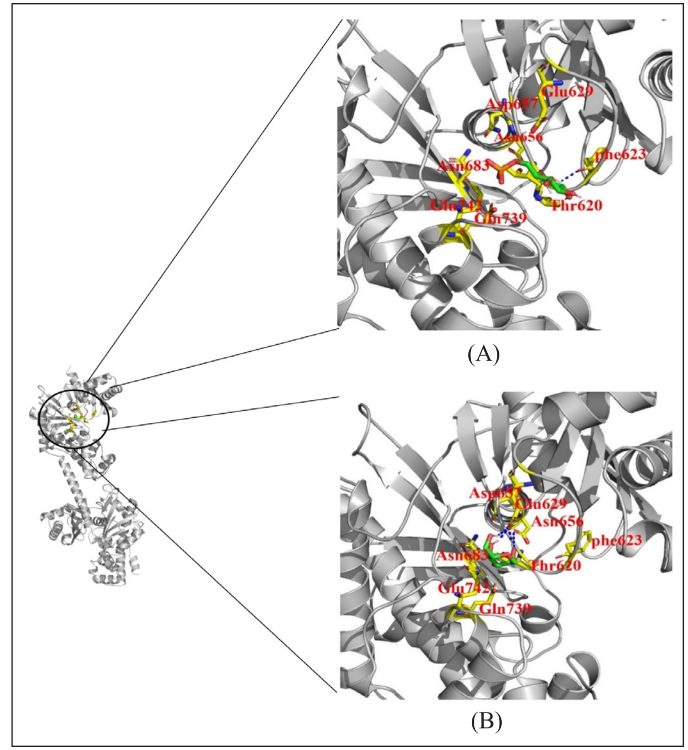


Figure 4. The interactions between BG6 and compound 14 with the residues involved in the catalytic site (A) HKII in complex with BG6 (green, red and orange color indicates carbon, oxygen and phosphorus atom of BG6), showing amino acid residues involved in Chain B with one H bond with Phe623. (B) HKII in complex with compound 14 (green and red color indicates carbon and oxygen atom respectively of compound 14), showing amino acid residues involved in Chain B with four H bonds (three formed with Glu629 and one with Asn656). Yellow and blue area indicate the carbon and nitrogen atoms of the particular amino acid residues, surrounding the ligand, while the blue lines indicate H bond.

Toxicity Test

Toxicity is proven essential in examining compounds for potential inclusion into drug development. The aim of toxicity studies is to ensure safety of the chemical compounds before they can be used as drugs or during clinical trials, as well as to determine the toxic effects of test substances (Arome & Chinedu, 2014). In several experiments, different types of toxicity tests (acute toxicity, sub-acute toxicity, and chronic toxicity studies) have been conducted to characterize the possible toxic effects of drugs that can range from minor to critical (Whitby et al., 2005; Ruiz et al., 2012). In this study, all compounds were assessed for toxicity using Toxicity Estimation Software Tool (TEST) and the results are reported in Table 7, 8, and 9, respectively. The TEST software utilizes the consensus method to determine the values of different compounds by using different tests end-point analysis approaches, where LD50 and LC50 values are exhibited for each compound (Tables 7-9). A smaller value indicates more toxic compound and vice versa. According to the data analysis results obtained in the present study, most of the compounds that were similar to 2DG were toxic, where only compound 4 (ZINC26898487) was non-toxic. In addition, the toxicity of the six top-ranked analogues of GLC showed that compound 9 (ZINC16930948) was found non-toxic, while the data for Compound 26 and 18 were not available (Table 8). Moreover, the top six analogues of BG6 were estimated for toxicity, and the best compound (compound 14, ZINC4403351) was confirmed as non-toxic. These toxicity results indicated that compound 4 (analogue of 2DG), compound 9 (analogue of GLC) and compound 14 (analogue of BG6) were non-toxic, and could be used in further *in vitro* HK2 inhibition studies. It should also be noted that all examined compounds were negative for mutagenicity. As already mentioned previously, 2DG is a known inhibitor of

Table 7
Estimation of toxicity values for the top six compounds similar to 2DG

Toxicity endpoint	Predicted value					
	Compound 4	Compound 18	Compound 6	Compound 8	Compound 19	Compound 21
Fathead minnow LC_{50}	606.12	N/A	609.12	535.21	404.59	N/A
Daphnia magna LC_{50}	45.87	N/A	45.87	1545.73	18.42	N/A
T.pyriformis IGC_{50}	1372.67	N/A	N/A	12725.59	N/A	N/A
Oral rat LD_{50}	1499.81	N/A	53.60	198.04	2496.78	N/A
Bioaccumulation factor	0.85	N/A	0.85	1.12	0.39	N/A
Developmental Toxicity	**Non-toxic**	N/A	Toxic	Toxic	Toxic	N/A
Mutagenicity	Negative	N/A	Negative	Negative	Negative	N/A

HKII, but it has shown toxic effect in *in-vitro* analysis (Fontaine et al., 2015), which gives an urge to search for compounds that can be developed into potent HKII inhibitor, which are non-toxic and safe to be administered.

Table 8
Estimation of toxicity values for the top six compounds similar to GLC

Toxicity endpoint	Predicted value					
	Compound 23	Compound 3	Compound 2	Compound 21	Compound 12	Compound 14
Fathead minnow LC_{50}	31166.04	2071.30	32728.60	8441.25	18386.86	32728.60
Daphnia magna LC_{50}	15038.30	474.16	9881.54	11182.63	14457.48	9881.54
T.pyriformis IGC_{50}	99007.65	N/A	34586.97	N/A	8704.83	34586.97
Oral rat LD_{50}	20804.48	2920.90	14201.45	13792.45	21288.21	14201.99
Bioaccumulation factor	0.31	0.41	0.28	N/A	0.22	0.28
Developmental Toxicity	Non-toxic	Non-toxic	Non-toxic	Non-toxic	Non-toxic	**Non-toxic**
Mutagenicity	Negative	Negative	Negative	Negative	Negative	Negative

Table 9
Estimation of toxicity values for the top six compounds similar to BG6

Toxicity endpoint	Predicted value					
	Compound 8	Compound 9	Compound 25	Compound 26	Compound 18	Compound 12
Fathead minnow LC_{50}	12144.34	12144.34	40759.57	N/A	N/A	40759.57
Daphnia magna LC_{50}	2350.97	2350.97	3596.98	N/A	N/A	3596.98
T.pyriformis IGC_{50}	6446.10	6446.10	15497.48	N/A	N/A	15497.48
Oral rat LD_{50}	9716.11	9716.11	12710.41	N/A	N/A	12710.41
Bioaccumulation factor	0.25	0.25	0.38	N/A	N/A	0.38
Developmental Toxicity	Non-toxic	**Non-toxic**	Toxic	N/A	N/A	Non-toxic
Mutagenicity	Negative	Negative	Negative	N/A	N/A	Negative

CONCLUSION

In conclusion, compounds similar to 2DG, GLC, and BG6 had been successfully identified by using ligand-based screening program. Six top-ranked compounds were then obtained for each query molecule, following the docking study, based on their binding energy and number of H bonds. Compound 9, which is an analogue of GLC, emerged as a good inhibitor, due to its best binding energy (-8.50 kcal/mol), and six H bonds (four of which are formed with Thr620, Asn683,cAsp657 and Glu742, respectively, and the remaining two with Asn656), as well as its predicted non-toxicity. Compound 14, which was most similar to BG6, had also been suggested as a potential inhibitor, because of its good binding energy (-7.09 kcal/mol), four H bonds formed with Glu629 and Asn656, negative mutagenicity, and its predicted non-toxicity. Finally, Compound 4, an analogue of 2DG, was suggested as a potential inhibitor of HKII, based on its good binding energy (-7.63 kcal/mol), with two H bond formed with Ser893 and Gln739, as well as its non-toxicity. The limitation of this studies is the utilization of only one compound library for screening the analogues of the query molecules, and through molecular docking, the interaction of the ligand and protein is analyzed in rigid condition, which might not be fully accurate. It was thus suggested that further molecular dynamics simulation studies could be conducted, where this analysis strengthens the choice of compounds to be utilized in the subsequent analysis. The selected compounds from this studies will be tested with purified recombinant HKII for inhibition analysis and subsequently will be tested against human fibroblast cells infected with DENV-2, to observe their effects in depleting DENV replication. In sum, Compound 4, Compound 9, and Compound 14 have great potentials as potent inhibitors of HKII, based on virtual screening and toxicity studies, and should be further tested in the future *in vivo* and *in vitro* inhibition studies.

ACKNOWLEDGMENT

The authors would like to thank the Ministry of Education Malaysia (MOE) for supporting this study through FRGS/1/2016/STG04/UIAM/02/1.

REFERENCES

Arome, D., & Chinedu, E. (2014). The importance of toxicity testing. *Journal of Pharmaceutical and BioScience, 4*(13), 146-48.

Basavannacharya, C., & Vasudevan, S. G. (2014). Suramin inhibits helicase activity of NS3 protein of dengue virus in a fluorescence-based high throughput assay format. *Biochemical and Biophysical Research Communication, 453*(3), 539-544.

Behnam, M. A. M., Nitsche, C., Vechi, M. S., & Klein, D. C. (2014). C-terminal residue optimization and fragment merging discovery of a potent peptide-hybrid inhibitior of dengue protease. *American Chemical Society, 5*(9), 1037-1042.

Bianchi, M., Rosano, C., Sabini, E., Rizzi, M., Deriu, D., Murshudov, G., … & Bolognesi, M. (1999). Binding of non-catalytic ATP to human hexokinase I highlights the structural components for enzyme-membrane association control. *Structure, 7*(11), 1427-1437.

Byrd, C. M., Dai, D., Grosenbach, D. W., Berhan, A., Jones, K. F., Cardwell, K. B., … & Jordan, R. (2018). A novel inhibitor of dengue virus replication that targets the capsid protein. *American Society for Microbiology Journals, 57*(1), 15-25.

Cosconati, S., Forli, S., Perryman, A. L., Harris, R., Goodsell, D. S., & Olson, A. J. (2010).Virtual screening with AutoDock: theory and practice. *Expert Opinion on Drug Discovery, 5*(6), 597-607.

Diamond, M. S., & Pierson, T. C. (2015). Molecular insight into dengue virus pathogenesis and its implications for disease control. *Cell, 162*(3), 488-492.

Fontaine, K. A., Camarda, R., & Lagunoff, M. (2014). Vaccinia virus require glutamine but glucose for efficient replication. *Journal of Virology, 88*(8), 4366-4374.

Fontaine, K. A., Sanchez, E. L., Camarda, R., & Lagunoff, M. (2015). Dengue virus induces and requires glycolysis for optimal replication. *Journal of Virology, 89*(4), 2358-66.

Fried, J. R., Gibbons, R. V., Kalayanarooj, S., Thomas, S. J., Srikiatkhachorn, A., Yoon, I. K., … & Cummings, A. T. (2010). Serotype-specific differences in the risk of dengue hemorrhagic fever: An analysis of data collected in Bangkok, Thailand from 1994 to 2006. *PLoS Neglected Tropical Diseases, 4*(3), 1-6.

Ganapathy-Kanniappan, S., & Geschwind, J. H. (2013). Tumor glycolysis as a target for cancer therapy: Progress and prospects. *Molecular Cancer, 12*(1), 1-11.

Gao, Q., Yang, L., & Zhu, Y. (2010). Pharmacophore based drug design approach as a practical process in drug discovery. *Current Computer-Aided Drug Design, 6*(1), 37-49.

Gershon, R., Crowther, A. J., Tikunov, A., Garcia, I., Annis, R., Yuan, H., … & Deshmuk, M. (2013). Hexokinase-2- mediated aerobic glycolysis is integral to cerebellar neurogenesis and pathogenesis of medulloblastoma. *Cancer and Metabolism, 1*(1), 2-17.

Godói, I. P., Lovato, L., Lemo, P., Araújo, V. E., Bonoto, B. C., & Godman, J. G. (2008). CYD-TDV dengue vaccine: Systematic review and meta-analysis of efficacy, immunogenicity and safety. *Journal of Comparative Effectiveness Research, 6*(2), 165-80.

Halim, S. A., Khan, H., Khan, A., Wadood, A., Mabood, F., Hussain, J., & Al-Harrasi, A. (2017). Targeting dengue virus NS-3 helicase by ligand based pharmacophore modeling and structural based virtual screening. *Frontiers in Chemistry, 5*(88), 1-6.

Idrees, M., Hussain, W., Rehman, H., Tayyab, G. N., Afzal, S., Fatima, Z., … & Waqar, M. (2012). Dengue virus serotypes 2 (DEN-2): the causative agent of 2011-dengue epidemic in Pakistan. *American Journal of Biomedical Science, 4*(4), 307-315.

Kaushik, N., Lee, S. J., Choi, T. G., Bai, K. Y., Uhm, H. S., Kim, C. H., … & Choi, H. E. (2015). Non-thermal plasma with 2-deoxy-D-glucose synergistically induces cell death by targeting glycolysis in blood cancer cells. *Scientific Reports, 5*(1), 1-11.

Kouretova, J., Hammamy, Z. M., Epp, A., Hardes, K., Kallis, S., Zhang, L., … & Steinmetzer, T. (2017). Effects of NS2B-NS3 protease and furin inhibition on West Nile and Dengue virus replication. *Journal of Enzyme Inhibition and Medical Chemistry, 32*(91), 712-721.

Kroemer, R. T. (2007). Structure-based drug design: docking and scoring. *Current Protein and Peptide Science, 8*(4), 312-328.

Lavecchia, A., & Giovannni, C. D. (2013). Virtual screening in drug discovery: A critical review. *Current Medicinal Chemistry, 20*(5), 60-71.

Lim, S. V., Basyaruddin, M., Rahman, A., & Tejo, B. A. (2011). Structure-based and ligand-based virtual screening of novel methyltransferase inhibitors. *Asia Pacific Bioinformatics, 12*(13), 13-24.

Lionta, E., Spyrou, G., Vassilatis, D. K., & Cournia, Z. (2014). Structure-based virtual screening for drug discovery: principles applications and recent advances. *Current Topics in Medicinal Chemistry, 14*(16), 1923-1938.

Lipinski, C. A., Lombardo, F., Dominy, B. W., & Feeney, P. J. (1997). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development setting. *Advanced Drug Delivery Reviews, 23*(1), 3-25.

Lunt, S. Y., & Heiden, M. G. V. (2011). Aerobic glycolysis: Meeting the Metabolic Requirements of cell proliferation. *Annual Review of Cell and Developmental Biology, 27*(1), 441-464.

Mastrangelo, E., Pezzullo, M., De Burghgraeve, T., Kaptein, S., Pastorino, B., Dallmeier, K., ... & Bolognesi, M. (2012). Ivermectin is a potent inhibitor of flavivirus replication specifically targeting NS3 helicase activity: new prospects for an old drug. *Journal of Antimicrobial Chemotherapy, 67*(8), 1884-1894.

Mirza, S. B., Salmas, R. K., Fatmi, M. Q., & Durdagi, S. (2016). Virtual screening of eighteen million compounds against dengue virus: Combined molecular docking and molecular dynamics simulation study. *Journal of Molecular Graphics and Modelling, 66*(2), 99-107.

Mishra, S., Agrahari, K., & Shah, D. K. (2017). Prevention and control of dengue by diet therapy. *International Journal of Mosquito Research, 4*(1), 13-18.

Morris, G. M., Goodsell, D. S., Pique, M. E., Lindstrom, W. L., Huey, R., Forli, S., ... & Olson, A. J. (2010). *Automated docking of flexible ligands to flexible receptors (autodock version 4.2.)*. Retrieved November 05, 2012, from http://autodock.scripps.edu/faqs-help/tutorial .

Mudin, R. N. (2015). Dengue incidence and the prevention and control program in Malaysia. *International Medical Journal Malaysia, 14*(1)*,* 5-9.

Munger, J., Bajad, S. U., Coller, H. A., Shenk, T., & Rabinowitz, J. D. (2006). Dynamics of the cellular metabolome during human cytomegalovirusinfection. *PLoS Pathogens, 2*(12), 1165-1175.

Nissink, J .W. M., Murray, C., Hartshorn, M., Verdonk, M. L., Cole, J. C., & Taylor, R. (2002). A new test set for validating predictions of protein-ligand interaction. *Structure, Function and Genomics, 49*(4), 457-471.

Parida, P., & Yadav, N. S. (2014). Structure based drug design from panduratin a against NS2B/NS3 protein. *Indo American Journal of Pharmaceutical Research, 4*(12), 2231-6876.

Patil, S. P., Ballester, P. J., & Kerezsi, C. R. (2014). Prospective virtual screening for novel p53–MDM2 inhibitors using ultrafast shape recognition. *Springer International Switzerland, 28*(2), 89-97.

Pelliccia, S., Wu, Y. H., Coluccia, A., Regina, G. L., Tseng, C. K., Famiglini, V., … & Silvestri, R. (2017). Inhibition of dengue virus replication by novel inhibitors of RNA-dependent RNA polymerase and protease activities. *Journal of Enzyme Inhibition and Medicinal Chemistry, 32*(1), 1091-1101.

Printz, R. L., Osawa, H., Ardehali, H., Koch, S., & Granner, D. K. (1997). Hexokinase II gene: structure, regulation and promoter organization. *Biochemical Society Transactions, 25*(1), 107-112.

Reddy, S. A., Priyadarshini, P. S., Kumar, P. P., Pradeep, H. N., & Sastry, N. G. (2007). Virtual protein and peptide science. *Current Protein and Peptide Science, 8*(4), 329-351.

Roe, B., Kensicki, E., Mohney, R., & Hall, W. W. (2011). Metabolomic profile ofhepatitis C virus-infected hepatocytes. *PLoS Pathoges, 6*(8), 1-8.

Ruiz, P., Begluitti, G., Wheeler, J., & Mumtaz, M. (2012). Prediction of acute mammalian toxicity using QSAR method: A case study of sulfur mustard and its breakdown product. *Journal of Molecules, 17*(8), 8982-9001.

Savic, L. J., Chapiro, J., Duwe, G., & Geschwind, J. (2016). Targeting glucose metabolism in cancer: A new class of agents for loco-regional and systemic therapy of liver cancer and beyond. *Hepatic Oncology, 3*(1), 19-28.

Schreyer, A. M., & Blundell, T. (2012). USRCAT: Real-time ultrafast shape recognition with pharmacophoric constraints*. Journal of Cheminformatics, 4*(11), 1-15.

Sing, J., Kumar, M., Mansuri, R., Sahoo, C. G., & Deep, A. (2015). Inhibitor desingning, virtual screening and docking studies for methyltransferase: A potential target against dengue virus. *Journal of Pharmacy and Bioallied Science, 34*(1), 2-8.

Vinodkumar, C. S., Kalapannavar, N. K., Basavarajappa, K. G., Sanjay, D., Gowli, C., Nadig, G., & Prasad, B. S. (2013). Episode of coexisting infections with multiple dengue virus serotypes in central Karnataka, India. *Journal of Infection and Public Health, 6*(4), 302-306.

Vyas, V., Jain, A., Jain, A. V., & Gupta, A. (2008). Virtual screening: A fast tool for drug design. *Scientia Pharmaceutica, 76*(3), 333-360.

Wang, Q. Y., Kondreddi, R. R., Xie, X., Rao, R., Nilar, S., Xu, H. Y., ... & Lakshminarayana, S. B. (2011). A translation inhibitor that suppresses dengue virus *in vitro* and *in vivo*. *American Society for Microbiology, 55*(9), 4072-4080.

Whitby, K., Pierson, T. C., Geiss, B., Lane, K., Engle, M., Zhou, Y., … & Diamond, M. S. (2005). Castanospermine a potent inhibitor of dengue virus infection In vitro and In vivo. *American Society for Microbilogy, 79*(14), 8698-8706.

Wichapong, K., Nueangaudom, A., Painwanit, S., Sippl, W., & Kokpol, S. (2013). Identification of potential hit compounds for dengue virus NS2B/NS3 protease inhibitors by combining virtual screening and binding energy calculation. *Journal of Tropical Biomedicine, 30*(3), 388-408.

Wick, A. N., Durury, D. R., Nakada, H. I., & Wolfe, J. B. (1956). Localization of primary metabolic block produced by 2-Doxygulose. *Journal of Biological Chemistry, 224*(2), 963-969.

Wijewickrama, A., Fernando, S., Geethal, S., Jayerathne, B., Perera, P. A., Abeynaike, S. A., … & Jeewandara, C. (2018). Emergence of a dengue virus serotype 2 causing the largest ever dengue epidemic in Sri Lanka. *American Journal of Biomedical Science, 4*(4)*,* 307-315.

World Health Organization. (2018, July). *Immunization, Vaccines and Biologicals*. Retrieved July 07, 2018, from http://www.who.int/immunization/diseases/dengue/en/.

World Health Organization. (2019, April). *Western Pacific Region: Dengue Situation Update Number 567*. Retrieved April 25, 2019, from https://www.who.int/docs/default-source/wpro.

World Mosquito Program. (2018, August). *First Fact about Mosquito Borne Disease*. Retrieved August 20, 2019, from http:// www.eliminatedengue.com/our-research/dengue-fever.

Yang, C. C., Hu, H. S., Wu, R. H., Wu, S. H., Lee, S. J., Jiaang, W. T., ... & Yueh, A. (2013). A novel dengue virus inhibitors BP13944, discovered by High-throughput screening with dengue virus replicon cells selectes for resistance in the viral NS2B/NS3 protease. *American Society of Microbiology, 58*(1), 110-119.

Yu, Y., Clippinger, A. J., & Alwine, J. C. (2011). Viral effects on metabolism: changes in glucose and glutamine utilization during human cytomegalovirus infection. *National Institute of Health, 19*(7), 360-367.

Yung, C. F., Lee, K. S., Thein, T. L., Leo, Y. S., Yung, C. F., Tan, K. L., … & Leo, Y. S. (2015). Dengue serotype-specific differences in clinical manifestation, laboratory parameters and risk of severe disease in adults, Singapore. *The American Society of Tropical Medicine and Hygiene, 92*(5), 999-1005.

Zhang, D., Li, J., Wang, F., Hu, J., Wang, S., & Sun, Y. (2014). 2-Deoxy-D-glucose targeting of glucose metabolism in cancer cells as a potential therapy. *Cancer Letters, 355*(2), 176-183.

Pertanika J. Sci. & Technol. 27 (4): 1649 - 1669 (2019)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Review article

Feature Extraction and Classification Techniques of MODI Script Character Recognition

Solley Joseph[1,2]* and Jossy George[1]

[1]*Department of Computer Science Christ (Deemed to be University), Bangalore, Karnataka, 560029 India*
[2]*Carmel College of Arts, Science and Commerce for Women, Goa, 403713 India*

ABSTRACT

Machine simulation of human reading has caught the attention of computer science researchers since the introduction of digital computers. Character recognition is the process of recognizing either printed or handwritten text from document images and converting it into machine-readable form. Character recognition is successfully implemented for various foreign language scripts like English, Chinese and Latin. In the case of Indian language scripts, the character recognition process is comparatively difficult due to the complex nature of scripts. MODI script - an ancient Indian script, is the shorthand form for the Devanagari script in which Marathi was written. Though at present, the script is not used officially, it has historical importance. MODI character recognition is a very complex task due to its variations in the writing style of individuals, shape similarity of characters and the absence of word stopping symbol in documents. The advances in various machine learning techniques have greatly contributed to the success of various character recognition processes. The proposed work provides an overview of various feature extraction and classification techniques used in the recognition of MODI script till date and also provides evaluation and comparison of these techniques.

Keywords: Classification techniques, feature extraction techniques, handwritten character recognition, MODI script OCR, offline character recognition survey

ARTICLE INFO

Article history:
Received: 31 August 2018
Accepted: 23 April 2019
Published: 21 October 2019

E-mail addresses:
solley.joseph@res.christuniversity.in (Solley Joseph)
frjossy@cimlavasa.in (Jossy George)
* Corresponding author

INTRODUCTION

In recent years a lot of work is done towards the digitization of physical documents as it facilitates better searches, modification and quicker addition. And this brings a need for software which automates the process of

ISSN: 0128-7680
e-ISSN: 2231-8526

extracting, analyzing, recognizing and storing the information from physical documents. Automatic recognition of either printed or handwritten text from document images and the conversion of it into machine-readable form is called as Optical Character Recognition (OCR) (Chaudhuri et al., 2017). Over the years researchers have worked in this area and have contributed in developing efficient character recognition for various languages (Jayadevan et al., 2011; Capin & Pulli, 1990). Indian language scripts are generally complex in nature, and hence recognition system for Indian scripts is comparatively difficult. Though technical challenges and lack of commercial market makes it difficult for OCR related research for Indian language, it is the need of the hour (Jayadevan et al., 2011). The character recognition process can be categorized into offline character recognition and online character recognition. In the case of the online method, the recognition process is done at the time of writing the character. In offline methods, the recognition is done on the scanned image of the character (Capin & Pulli, 1990). The focus of this work is offline character recognition of MODI script.

The process of character recognition involves five major steps such as Pre-processing, Segmentation, Feature Extraction, Classification and Post-processing. A detailed discussion of each step is given in section 2.

Historical Importance of MODI Script

MODI script was developed in Devagiri in the 12th century. Hemadri Pandit also known as Hemadpant, a famous leader of Yadav dynasty, introduced MODI script. He was inspired by the cursive Sinhalese script and was instrumental in developing a similar one for Marathi which involves less lifting of hand while writing and which can be written very fast (Ramraje, 2013). MODI was easy to write and was commonly used as an official script for writing Marathi until 1950, and there are vast amount of documents preserved in various libraries (Khillari, 2008; Thakre, 2016). Because of the difficulty in printing MODI script, its usage was stopped, and Devanagari was used for writing Marathi. In addition to Marathi, there are other languages like Urdu, Kannada, Gujarati, Rajasthani, Hindi, and Tamil which used MODI script for writing (Pandey, 2011). The bulk of documents and correspondence from before Chhatrapati Shivaji's times are written in MODI Script. Large volumes of MODI documents are preserved in India as well as other Asian and European countries. Bharat Itihas Sanshodhan Mandal in Pune, Tanjavur's Saraswati Mahal, Rajwade Sanshodhan Mandal and Dhule (Maharashtra) known to have large collections of such documents (Khillari, 2008). The Pune division of the State Archive Department also has a very valuable collection of oldest and rare manuscripts dating back to the Peshwa dynasty and times of Chhatrapati Shivaji (Deshpande, 2013). Goa's Directorate of Archives and Archaeology also have a collection of MODI documents as land and revenue records of various talukas of the state. These are loose manuscripts which are archived in the form of books. The archive has 13,000 pages in 113 books of one taluka itself and other loose manuscripts which are yet to be documented. It is observed that some of the MODI documents which are archived in various places are on the verge of degradation due to improper facilities. Digitization of these documents will help to preserve this large repository of knowledge. Realizing the fact, that it is impossible to study Maratha history without knowing MODI script, researchers and enthusiasts work towards keeping the script alive (Times, 2014; Savant & Rakshana, 2005). It has been observed that there is strong effort to revive MODI script by publishing articles written in the script. In addition, several institutions now offer tutorials for learning MODI. Workshops and formal courses conducted by the Maharashtra State Department of Archives and Bharat Itihas Samshodhak Mandal (BISM) in Pune are examples of such courses. In parallel with the revival of the script the effort to catalogue and manage large volumes of MODI documents stored at various libraries are also taking place. Script Encoding Initiative (SEI) of the University of California, Berkeley, has initiated and completed a project to encode the script, and currently is supported in Unicode (Berkeley-Linguistics, 2018). The Government of India has initiated the work of cataloguing of these manuscripts by appointing trained MODI experts for the job (Pandey, 2011). Aimed at preserving and propagating MODI script, CDAC (Centre for Development of Advanced Computing) Mumbai has done a considerable amount of work on MODI script and has developed MODI script software. CDAC made a website enabling users to convert text from Devanagari script to MODI and *vice-a-versa* (CDAC, 2016). In a project funded by the Government of India, Tamil University has taken steps to digitize and catalogue MODI documents by converting them into a Portable Document Format (Chen & Wang, 2000). In the year 2013, the Govt. of Maharashtra allocated 80 Lakhs for the digitization of MODI manuscripts. The reports state that there is a serious effort to keep MODI scripts alive (Hindu, 2007, 2013). In addition to that individual researchers have carried out researches in MODI script character recognition using various methodologies.

Description of MODI Script

MODI script is an ancient script which was used in the 12th Century for writing Marathi. This script was widely used between the 12th century and 20th century. The name 'MODI' believed to be derived from the Marathi verb moḍaṇe (Marathi: मोडणे), which means "to bend or break" (Pandey, 2011). The script was written by 'Boru' or 'lekhan'. 'Lekhan' was a pen created with the 'Bamboo'. MODI script is derived from the Nagari family of scripts and intended for continuous writing. Although MODI is based upon Devanagari, there are some significant differences. The differences are evident in letter forms, rendering behaviors, and orthography of both the characters. The behaviors of these characters in certain environments, such as consonant-vowel combinations and in consonant conjuncts that are standard features of MODI orthography and is different from Devanagari script. The MODI script has 46 distinctive letters, of which 36 are consonants and 10 are vowels (Figure 1). MODI was written as characters hanging on a horizontal line which are drawn across the page. Termination symbol for sentence or word was not used in MODI script and hence word segmentation is a very difficult task in the case of MODI script documents. The speed of writing was increased due to the elimination of these symbols as it avoided the lifting of the "Boru' too often (Kulkarni et al., 2015a).

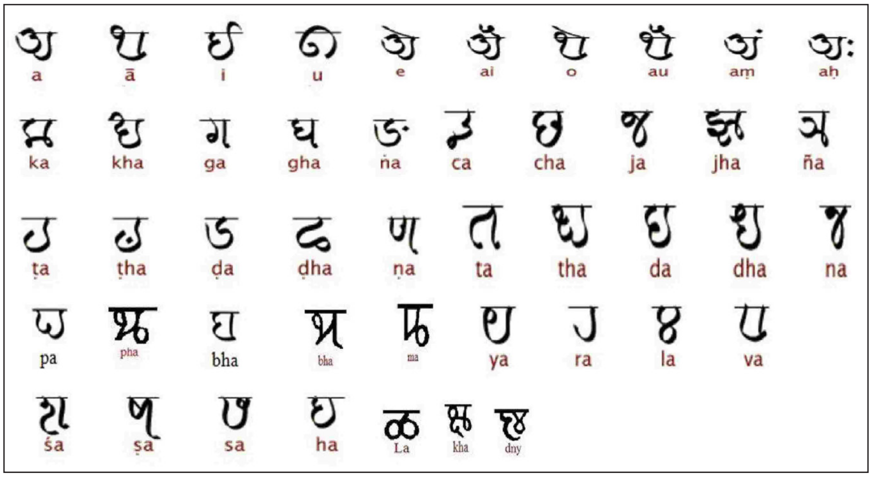


Figure 1. Basic character set of MODI script

The review of literature reveals that there are several styles of the MODI script based on the era. The different styles of the script are as follows:- Adya Kalin (proto-MODI) of 12th century, Yadav Kalin of 13th century (Yadav dynasty), Bahamanikalin which was during 14th to 16th century, Sivakalin of 17th century, Chitnisi of 18th century, Peshvekalin which lasted till 1818, and Anglakalin which was popular during 1818 to 1952 (Figure 2). Among the various styles, Bahamanikalin, Chitnisi, Peshvekalin and Anglakalin are found to be commonly used in historical manuscripts (Pandey, 2011). The Anglakalin is on the messier side while the Peshvekalin has a more aesthetic flow. In addition, a style of MODI was used in the primary school books (Figure 2) produced during the 19th and 20th centuries. Interaction with the experts of MODI scripts reveals that in the Bahamanikalin MODI documents there is a tendency of Parsi word usage.

In this paper, an effort is carried out to review various feature extraction and classification techniques used in the recognition of MODI Script till date (August 2018), which will promote MODI script related research. The remaining part of the paper is organized into four sections. Section 2 deals with the various stages of the handwritten character recognition process and the commonly used methods for the same. Section 3 deals with a detail discussion on MODI character recognition with an emphasis on various feature extraction as well as classification techniques used in MODI character recognition. Section 4 deals with current trends in character recognition. Section 5 gives the conclusion of the paper.

Figure 2. Various styles of MODI script

HANDWRITTEN CHARACTER RECOGNITION PROCESS

Handwritten character recognition involves Pre-processing, Segmentation, Feature Extraction, Classification and Post-processing as shown in Figure 3.

Pre-processing is used to eliminate the noise of the document. There are various pre-processing techniques such as Noise Reduction, Filtering, Morphological operations, Skew detection and Normalization can be used at this stage (Arica & Yarman-Vural, 2001). Skew detection, skew correction, and shirorekha extraction was performed on MODI documents and achieved an accuracy of more than 90% (Tamhankar & Kolhe, 2018).

In the segmentation step, the document is subdivided into components like words, lines, and characters. The accuracy of this stage affects the accuracy of the entire process. An experiment for line segmentation of MODI documents yielded 91% accuracy in the segmentation of overlapping lines and 94% accuracy in extracting touching lines (Kulkarni et al., 2017). The segmentation of connected handwritten character is the main bottleneck

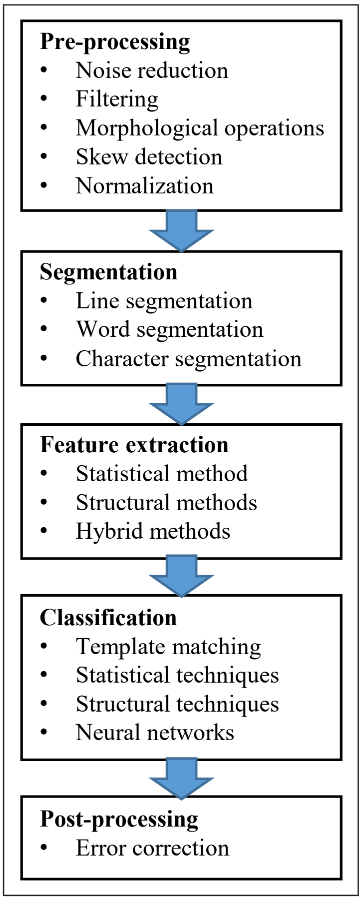


Figure 3. Various stages of the character recognition process

in the OCR system (Chen & Wang, 2000; Dhaka & Sharma, 2015). A performance analysis of character segmentation approach for cursive script recognition is carried out in which the geometric features of the character are analyzed for the segmentation (Rehman & Saba, 2011). An algorithm using top profile projections is efficiently implemented to segment touching characters in the upper zone of printed Gurmukhi script and has achieved a recognition rate of 91% (Jindal et al., 2009). A fuzzy multifactorial analysis is used for the segmentation of touching printed characters of Bangala and Devanagari (Garain & Chaudhur, 2001). A Pixel Plot and Trace and Re-plot and Re-trace (PPTRPRT) technique is used for segmentation of Devanagari handwritten scripts for extracting text regions (Dhaka & Sharma, 2015). Word segmentation in handwritten documents is a difficult task as the gap between words is not uniform. MODI script documents do not have any space between words as the characters are continuously written in the line. Though word segmentation is not been experimented in MODI Script, researchers have experimented word segmentation of for various other scripts.

The extraction of the most representative information from the raw data is called as feature extraction. A feature point is the distinguishing feature of the character image, and such features will be recognized at this stage, with a fixed number of feature variables (Chacko et al., 2012). The extracted features of each class are used for distinguishing it from the other classes. The feature extraction is the most important step in the character recognition process as it highly influences the recognition rate of the entire process. There are various feature extraction methods available for the selection of important features. Identifying the suitable feature extraction method is essential in achieving an excellent recognition performance. The feature extraction methods for character recognition can be broadly classified as statistical and structural methods. In the case of statistical feature extraction, the feature extraction is based on the statistical distribution of points which includes zoning, histograms and moments. Structural features are based on the appearance of the character or its geometrical shape like directions of the stroke and intersection of strokes etc. (Jayadevan et al., 2011). It is observed that neither the structural nor statistical information can represent a complex pattern alone (Goswami & Sharma, 2013). Therefore an approach integrating both statistical and structural features is also implemented by many researchers. Research in the field of character recognition indicates that the recognition accuracy depends heavily on the feature extraction methods and classification methods used in the recognition process. The commonly used feature extraction techniques used for MODI character recognition are based on Statistical methods such as Zernike moment, Hu's seven invariant moments, Chain Code Histogram and Intersection/Junction. The efficiency of the Zernike feature extraction method was tested for MODI Script character recognition by various researchers (Kulkarni et al., 2014, 2015b; Sadanand et al., 2015a, 2015b). The observations made by them indicate that the use of Zernike Complex moments features with zoning increased the performance. A similar method was used for Oriya alphabets (Tripathy, 2010), Devanagari (Marathi) Compound Character Recognition (Kale et al., 2014), and Arabic character (Oujaoura et al., 2012) and achieved good performance. The performance of Chain Code histogram and Insertion junction feature extraction methods was tested on MODI as well as Devanagari script (Chandure & Inamdar, 2016). Chain Code Histogram based feature extraction method is used for handwritten Bangla characters (Bhattacharya et al., 2006) and resulted in acceptable recognition performance. Feature extraction using AMIs (Affine Moments Invariants) technique has been proposed for MODI character recognition in order to increase the recognition accuracy and reliability (Patil, 2016). A hybrid approach for feature extraction which combines two techniques: Moment Invariant and Affine Moment Invariant is reported to have achieved an average recognition rate of 89.72% (Gharde & Ramteke, 2016). They claim that the hybrid technique also resulted in extracting minimum features. A similar approach is used for Farsi Character Recognition System as well (Alavipour & Broumandnia, 2014). Their research achieved an accuracy of 96.5%.

The classification stage is instrumental in mapping an unknown sample into a predefined class. The quality of the extracted features determines the performance of the classifier. Various methodologies of pattern recognition are used at this stage for the classification purpose. The techniques for Character Recognition can be categorized in four general approaches of pattern recognition as template matching, statistical techniques, structural techniques and neural networks (NNs) (Jain et al., 2000). This classification is not distinct as a technique in one approach may also be classified as a member of other (Arica & Yarman-Vural, 2001). There exist various methods in each of these approaches. An analysis of different classification methodologies used in OCR of various languages was carried out in which the author discusses various classification algorithms used for pattern recognition (Sharma, 2013). The author discussed methods such as Support Vector Machine, Artificial Neural Network, K Nearest Neighbor, Naïve Bayes Classifier, Decision Tree, Linear Discriminant Analysis (LDA), Quadratic Discriminant Analysis (QDA), and Maximum entropy classifier for pattern classification. An insight into different learning based classification methods such as SVM, Artificial Neural Network (ANN) and various statistical methods were discussed in Liu & Fujisawa, (2008). A survey on offline recognition of handwritten Devanagari scripts gives an insight into various algorithms used in offline handwritten recognition (Ramteke & Rane, 2012).

In the post-processing stage, the errors are corrected and the ambiguities of the OCR are resolved. Post processing is done at the word level, sentence level, and semantic level.

MODI SCRIPT CHARACTER RECOGNITION

Though handwritten character recognition is successfully implemented for various foreign language scripts, most of the Indian language scripts still need attention. The development of OCR for such script is difficult due to the complex nature of scripts (John et al., 2011). MODI being an ancient script and not in the list of scheduled official scripts of Indian languages, very limited work is done towards the handwritten character recognition of MODI script compared to other Indian scripts. Character recognition of MODI script is a difficult task due to the fact that characters in MODI are cursive, continuous, unconstrained, and the existence of highly similar shaped characters. Segmentation is another challenging task of MODI character recognition process. Major challenges are caused due to noise and degradation, the presence of multiple skews, fluctuation in illumination, uneven alignment, slanting lines, overlapping lines and touching lines (Gharde & Ramteke, 2016). Research in this field indicates that MODI character recognition needs a different approach/ method as compared to other Indian script character recognition. Unlike other scripts, MODI script document does not have any termination symbol for words or sentences, and therefore word segmentation techniques cannot be applied on MODI manuscripts.

Review of the literature indicates that only thirteen published works are available on the character recognition of MODI script which is discussed in the following section. In addition there are few studies which deal with the theoretical analysis, survey, and analysis of various algorithms used in MODI script, etc. A study carried out on reviewing the various feature extraction techniques conclude that feature extraction methods of MODI character recognition are still in the infancy stage compared to other Indian languages. The study states that segmentation is the most concerned part in the case of MODI manuscripts, especially as there are overlapping characters and lines in ancient MODI documents (Kulkarni et al., 2015a). An analysis of some of the existing systems on handwritten character recognition system of Indian Script with a focus on MODI script gives an insight to the various methods which can be used for MODI script recognition (Balbudhe et al., 2016). A theoretical analysis of MODI Script character recognition gives an insight into the basic features of the ancient script (Besekar, 2013). A comparison of the character recognition process of MODI, Devanagari, and Roman scripts was performed in this work and found that in the case of MODI script structural features were difficult to extract. This study reveals the complexities occurring in MODI script character recognition and also states that the recognition process is more difficult compared to other Indian languages. Cursive natures of the character, complex structure of some of the characters, inconsistency in writing style, shape similarity of characters, etc. are some of the factors which make the recognition task extremely difficult. A feature extraction technique using Affine Moment invariant has experimented with a fuzzy logic classifier in (Patil, 2016). This study stated that the classification got a higher accuracy rate if the deformation between the template and the unknown character was approximately affine. MODI character recognition and an attempt to convert it into corresponding English character are carried out in (Rathi, 2015). This study listed out different steps of the image processing which could be used for MODI script character recognition. A work towards skew detection, skew correction and segmentation of MODI document is carried out in which a novel algorithm is implemented for the shirorekha extraction from the image (Tamhankar & Kolhe, 2018). The experiment resulted in an accuracy of 93% in skew detection and in the case of shirorekha extraction the accuracy was 90.27%. In another experiment on skew detection of MODI document (Deshmukh et al., 2017), the authors experimented with a horizontal projection based approach, for the detection of skew angle in old MODI documents. They have also performed a comparison of the proposed algorithm with the benchmarking techniques illustrated by Mahnaz and Maher, (2015), and claim that the proposed method resulted in better accuracy, with an average success rate of 96.49% as compared to 95% in the benchmarking techniques. A study on various Thresholding Technique used in the ancient document images is discussed in (Jyoti & Kumar, 2016). The authors had evaluated fourteen different Thresholding Techniques which could be used for degraded image binarization and concluded that each technique had its own benefits and limitation.

MODI script character recognition involves all the five stages, as illustrated in the previous session. The focus of this paper is the feature extraction and classification techniques used by various researchers in the process of MODI script character recognition.

Feature Extraction Techniques used in MODI Script Character Recognition

In MODI script character recognition both structural, statistical as well as a hybrid method of feature extraction are used by various researchers. The most commonly used are Statistical based methods

Zernike and Zernike Complex Moments. The moment-based approach is found to be more commonly used in the recognition of MODI characters as well as numerals and found to have achieved better accuracy compared to the rest of the methods used. The efficiency of the Zernike feature extraction method was tested for MODI Script character recognition by various researchers (Kulkarni et al., 2015a; Sadanand et al., 2015a, 2015b) The observations made by them indicated that the use of Zernike Complex moments features with zoning increased the performance. Zernike complex moments in combination with Decision Tree was found to achieve an accuracy of 97.68 % (Kulkarni et al., 2016). A similar method was used with different zoning patterns (Sadanand et al., 2015a) and Euclidian distance classifier and achieved an accuracy of 94.92 (using Zernike moments), and 94.78 (using Zernike complex moments). Every character was divided into six zoning patterns with thirty-seven zones. The zoning patterns were created using geometrical shapes. Hu's Seven Moments and Zernike was used for the recognition of MODI numerals in (Kulkarni et al., 2015a) and achieved an accuracy of 70% in the case of Hu 's seven and 86.6 in the case of Zernike moment. When the Zernike moment feature is used in combination with Euclidean distance classifier for the recognition of MODI character, the accuracy is 82.61% (Sadanand et al., 2015b). The observation made by them indicates that the Zernike moment performs more efficiently compared to Hu's 7 moments.

Chain Code based Approach. A recognition model based on image centroid and Chain code was used in MODI Script character recognition (Besekar, 2012). The accuracy achieved was 65.3% to 73.5% in combination with a two-layer Feed Forward Neural Network classifier in the case of MODI Script vowels. The same method was experimented in the recognition of the entire character set of MODI script as well as for Devanagari Script, in combination with BPNN, KNN and SVM. The accuracy achieved was 37.5%, 60% and 65% respectively (Chandure & Inamdar, 2016). It has been reported that Devanagari script dataset yielded better performance in the same experiment, with an accuracy of 70%, 87.5% and 87.5% respectively, in combination with BPNN, KNN and SVM. The observation from this experiment was that the misclassification of similar shaped vowels affected the performance in the case of MODI script.

Zone-based Approach. A zone-based feature extraction technique was used in Besekar and Ramteke (2012) for the recognition of MODI script Numerals. The feature set used in this experiment was created with the help of 4 equal square zone of size 15 x 15 and their polar co-ordinates, Variance, Theta, and Rh distance. The recognition accuracy achieved was 93.5%.

Structure Similarity Approach. A Structure Similarity Approach was experimented in combination (Ramteke & Katkar, 2013) for MODI Script character recognition. Image quality was measured using the Structural Similarity method. The classification method used in the experiment was Kohonen Neural Network & Back Propagation Neural Network (BPNN) and yielded a recognition accuracy of 91% to 97%.

Heuristic Approach. A feature extraction method using empirically determined heuristics is implemented for the character recognition of MODI script and resulted in an average recognition rate of 91.2% (Maurya & Maurya, 2018). In this case, a hybrid feature space was created using normalized chain code shape descriptions. It was reported that this hybrid approach had an advantage that allowed more flexible inputs such as the character of varying sizes. The experiment was done on thirty-three characters of MODI script with a data size of 3200 characters.

Hybrid Approach. A Hybrid approach was experimented by combining two methods in the recognition of MODI script (Gharde & Ramteke, 2016). Invariant and Affine invariant moments were combined for the feature extraction purpose. Eighteen features were extracted from a single numeral/character. The recognition accuracy achieved was 89.72 with SVM classifier. The highlight of this work was that with minimum features, high accuracy was achieved. The original manuscript of MODI script was used as the dataset in this case.

The Percentage of average accuracy of various Feature Extraction Techniques used in the recognition of MODI character data set is shown in Figure 4.

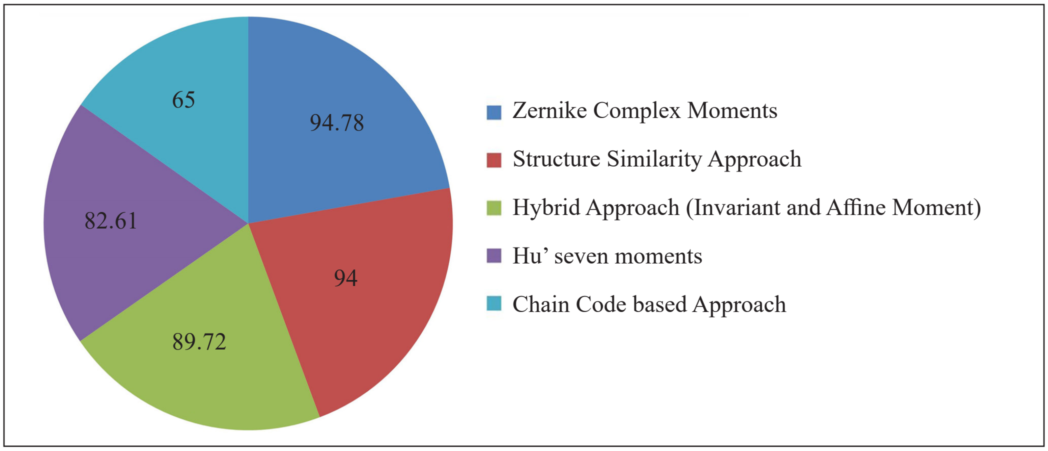


Figure 4. The performance (% of accuracy) of various feature extraction techniques used in MODI script character recognition

The Classification Techniques used in MODI Script Character Recognition

Decision Tree. There are various methods experimented by researchers for MODI script character recognition, within which, Decision Tree classifier was found to achieve the highest recognition rate. Zernike and Zernike complex moments in combination with Decision Tree was found to achieve an accuracy of 97.68% (Kulkarni et al., 2016).

Euclidean Distance Classifier. Euclidean distance classifier is the most commonly used method for MODI character recognition. Zernike and Zernike complex moments in combination with Euclidian distance classifier has achieved 94.78% (Sadanand et al., 2015a). In another experiment Euclidean distance classifier is implemented with the Hu's seven and Zernike features for the MODI numerals (Kulkarni et al., 2015a) as well as MODI characters (Sadanand et al., 2015b). In the case of MODI numerals the recognition accuracy was 86.6%, and in the case of MODI script character, the recognition rate was 82.6%.

Support Vector Machine. Support Vector Machine-based classification is implemented by two researchers for MODI script character recognition. In the first case (Chandure & Inamdar, 2016) the SVM classifier with chain code features achieved an accuracy of 65%, and with intersection junction feature the accuracy was 47.5 %. In the second case (Gharde & Ramteke, 2016) SVM classifier was used with Moment Invariant and Affine Moment Invariant feature and the accuracy was 89.72 %.

Back Propagation Neural Network. Back Propagation Neural Network based classification technique was used for the recognition of MODI characters in two experiments. It was used in combination with a feature extraction method of structure similarity and the accuracy rate was 91% to 97% (Ramteke & Katkar, 2013). In the second case it was implemented using Chain code histogram and achieved an accuracy of 37.5%, and with intersection junction feature the accuracy was only 15% (Chandure & Inamdar, 2016).

Kohonen Neural Network. Kohonen Neural Network based classification method was implemented by two researchers. With structural similarity feature extraction method the accuracy achieved was 91 to 97% (Ramteke & Katkar, 2013). In the second case, it was used in combination with Otsu's Binarization Algorithm for recognition of 22 MODI characters, and an overall character recognition rate of 72.6% was achieved (Anam, 2015).

Feedforward Neural Network. Feedforward Neural Network Classification technique was implemented on MODI vowels (Besekar, 2012) with chain code histogram based classification method and the recognition rate achieved was 65.3% to 73.5% with Zernike moments based features and the recognition rate achieved was 97.68%.

K-Nearest Neighbor Classifier. K-Nearest Neighbor classifier was implemented for MODI Character recognition (Chandure & Inamdar, 2016) in combination with Chain Code histogram based feature extraction and achieved an accuracy of 60%.

Variance Table. Variance Table was used for classification of MODI numerals (Besekar & Ramteke, 2012) and achieved a recognition rate of 93.5%. The feature set used for this experiment was created with the help of 4 equal square zone of size 15 x 15 and their polar co-ordinates, Variance, Theta, and Rh distance.

The average recognition accuracy of the various classification techniques used in the recognition of MODI character data set is shown in Figure 5.

The character recognition of MODI script implemented by the various researchers is consolidated in the Table 1. The data set for the experiment includes MODI Characters, MODI Numerals and a subset of MODI characters.

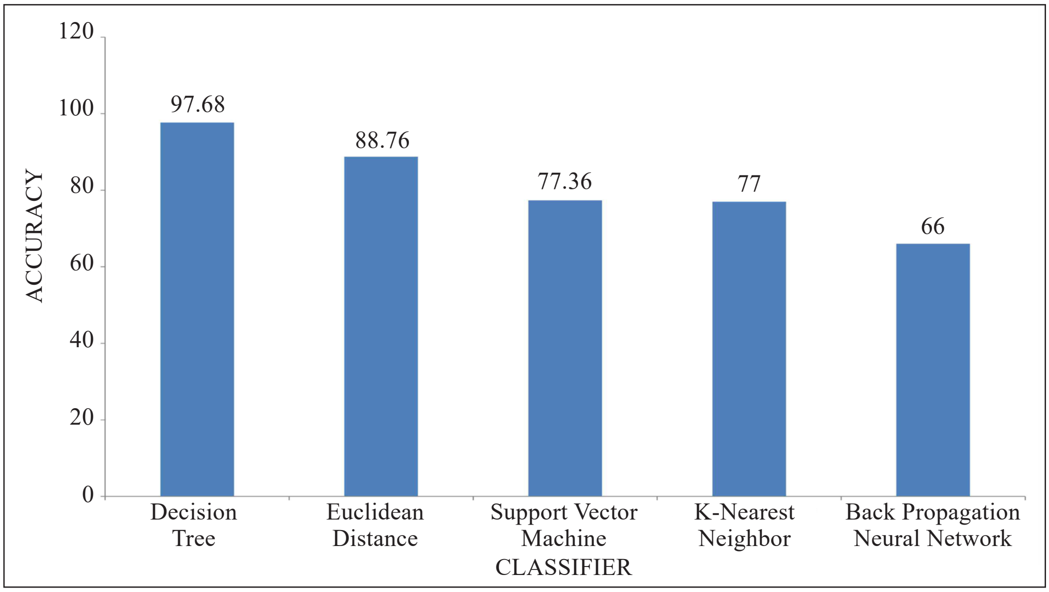


Figure 5. Average recognition accuracy (%) of various classifiers in MODI script character recognition

Table 1
List of methods used in MODI Script Recognition and comparison of the results achieved

Sr. No	Author	Feature Extraction	Classification	Data type	Accuracy
1	Kulkarni et al., 2016	Zoning, Zernike and complex Moments	Decision Tree	Characters	**97.68%**
2	Kulkarni et al., 2015a	Zoning, Zernike Moments and Zernike complex moments	Euclidian Distance	Characters -	**94.92% (Zernike) & 94.78% (Complex)**

Table 1 *(continue)*

Sr. No	Author	Feature Extraction	Classification	Data type	Accuracy
3	Ramteke & Katkar, 2013	Structure Similarity	KNN, BPNN	Characters	**91% to 97%**
4	Gharde & Ramteke, 2016	Hybrid–Combination of Invariant and Affine Moment Invariant	SVM Classifier	Characters	**89.72%**
5	Sadanand et al., 2015b	Hu's and Zernike moments	Euclidian Distance	Characters	82.61%
6	Chandure & Inamdar, 2016	Chain Code Histogram & Intersection/Junction features.	KNN, BPNN and SVM classifier	Characters	**KNN-60% BPNN-37% SVM-65%**
7	Maurya & Maurya, 2018	Chain Code	Empirically determined heuristics	33 Characters	**91.20%**
8	Anam, 2015	Chain Code	Kohonen Neural Network	22 characters	**72.60%**
9	Besekar, 2012	Chain code histogram and normalized chain code histogram	Feedforward NN	Vowels	**65.3% to 73.5%**
10	Besekar & Ramteke, 2012	Zoning - Variance of theta, Rh distance and centroid of the zone	Variance table	Numerals	**93.50%**
11	Kulkarni et al., 2014	Hu's and Zernike moments	Euclidian Distance	Numerals	**86.66%**
12	Kulkarni et al., 2015a	Hu's invariant features and Zernike moments	Euclidean' distance	Numerals	**70% & 86.6%**

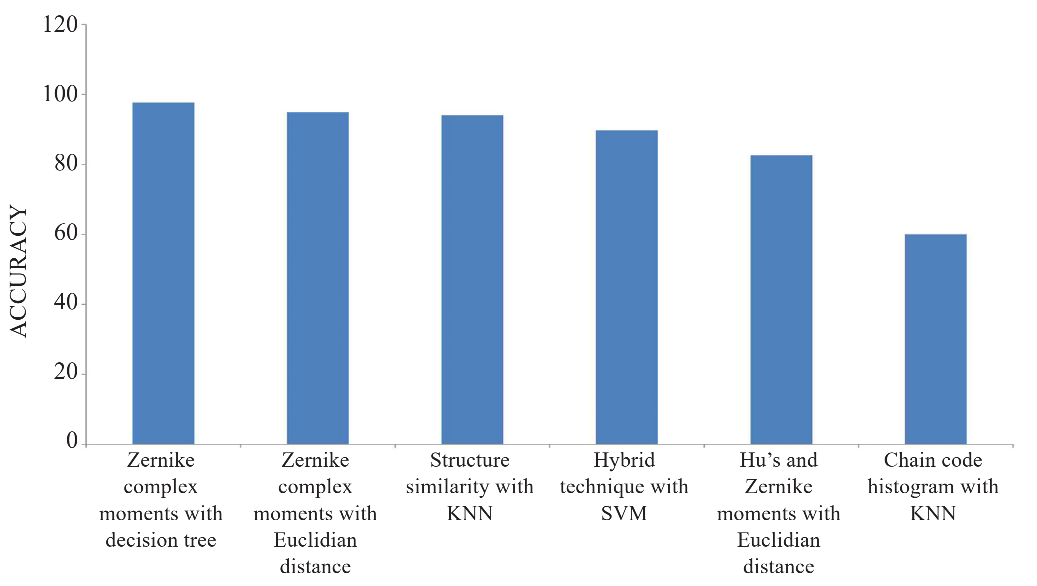


Figure 6. Various feature extraction techniques and classification techniques used in MODI script character recognition and their performance

The consolidated list of all the feature extraction and classification techniques used by various researchers in MODI character recognition is shown in Figure 6. The dataset used is forty-four MODI characters. The feature extraction techniques used by various researchers in MODI Script recognition are Zernike and Complex Moments, Structure Similarity, Hybrid Technique by combining Invariant and Affine Moment Invariant, Hu's and Zernike Moments and Chain Code Histogram. The classification techniques used are Decision Tree, Euclidian Distance, KNN and SVM.

CURRENT TRENDS IN CHARACTER RECOGNITION

Genetic Programming based Algorithms, Deep Learning based Algorithms, Wavelet Energy Feature (WEF) and Extreme Learning Machine (ELM) are effectively used in the recognition of various Indian languages. The GP algorithm is widely used in signal processing and pattern recognition (Jindal et al., 2009). Metric learning based Recognition Algorithm called Genetic Programming Metric Learning (GPML) is successfully implemented for Persian/Arabic character recognition (Sepahvand et al., 2017). It is learnt that this algorithm was efficient in extracting the differentiating geometric shape of the characters and thus producing a minimum set of features. A combination of Wavelet energy feature (WEF) and Extreme Learning Machine (ELM) is used effectively and yielded high recognition accuracy in handwritten recognition of Malayalam scripts (Chacko et al., 2012). The wavelet energy (WE) which is derived using wavelet transform is a new and robust parameter. Character recognition of Marathi compound characters is implemented using wavelet approximation features and modified wavelet features and neural network classifier (Shelke & Apte, 2011). In this study, it is observed that recognition accuracy is increased due to the modified wavelet feature extraction method. Deep learning which was first presented by Geoffrey Hinton et al (Hinton, 2007; Hinton et al., 2012) is gaining popularity in the recent years for various pattern recognition tasks and is efficiently implemented by various researchers (Bengio, 2009; Deng, 2014; Hinton & Salakhutdinov, 2006; Larochelle et al., 2007). The multiple layers in this kind of networks can efficiently handle more complicated functions than shallow ones.

Convolutional Neural Network (CNN) is another method which is gaining popularity for the feature extraction as it is one of best performing character recognition (Cires et al., 2011) (Maitra et al., 2015) algorithms. CNN is effectively used for handwritten characters (MNIST Dataset) (Cires et al., 2011), and Multi-script Numeral Dataset (Maitra et al., 2015) and yielded a high level of accuracy and is in the upfront of best-performing character recognition algorithms. The deep Convolutional Neural network is implemented for the character recognition of Hangul Script (Kim & Xie, 2015) and to improve upon the performance of the networks. A framework using Recurrent Neural Network (RNN), for the recognition and drawing of handwritten Chinese characters is reported to have achieved a new state-of-the-art performance. It is used for online character recognition as it can directly deal with the raw sequential data, unlike the CNN based method where some image-like representations are required (Zhang et al., 2018). RNN classifier is also used effectively for offline character recognition of Arabic and Latin scripts (Chherawala et al., 2016). A weighted vote combination of RNN classifiers each trained with a particular feature set is used in the experiment.

CONCLUSION

The primary objective of this work is to review the literature to identify the methodologies used by various researchers, in the character recognition of MODI script with an emphasis on the feature extraction and classification techniques. It is observed that MODI script character recognition is still in its initial stages. MODI being an ancient script and not in the list of scheduled official scripts of Indian languages, contributes to the fact that limited research work has been done towards it, compared to other Indian scripts. Research and development in MODI script are necessary to extract the information on MODI manuscripts which are stored in various parts of India and abroad. Though handwritten character recognition is successfully implemented for various foreign language scripts, most of the Indian language scripts still need attention. The development of OCR for such script is difficult due to the complex nature of scripts. Research in this field indicates that MODI character recognition needs a different approach/method as compared to other Indian script character recognition. Pattern recognition of MODI script is tedious because of various factors like the shape similarity of characters and inconsistency in writing style. Unlike other scripts, MODI script document does not have any termination symbol for words or sentences, and therefore word segmentation techniques cannot be applied on MODI manuscripts. Only a few works have been published so far and all the published work has been reviewed in this study. Segmentation is found to be a difficult stage in the case of ancient MODI documents and many of the published work of MODI script character recognition, segmentation methods used were unclear. In most of the work which is done on MODI script character recognition, the data set is specially generated in a controlled environment. But the documents which are available in various libraries are ornamental MODI script, and the recognition process of those documents is a difficult process. There is a need for more effort to unveil the historic information written in various MODI manuscripts. Implementation of more efficient methods at the various stages of the character recognition process will increase the performance of character recognition of MODI script. The purpose of this paper is to set a foundation for those who want to carry out research in this field.

ACKNOWLEDGEMENT

The facilities provided by the Department of Computer Science, Christ (Deemed to be University), Bangalore, India, and Carmel College of Arts, Science & Commerce for Women, Goa, India, for carrying out this study are gratefully acknowledged.

REFERENCES

Alavipour, F., & Broumandnia, A. (2014). Farsi character recognition using new hybrid feature extraction methods. *International Journal of Computer Science, Engineering and Information Technology (IJCSEIT)*, *4*(1), 15-25.

Anam, S. (2015). An approach for recognizing modi lipi using otsu's binarization algorithm and kohonen neural network. *International Journal of Computer Applications*, *111*(2), 28-34.

Arica, N., & Yarman-Vural, F. T. (2001). An overview of character recognition focussed on off-line handwritting. *IEEE Transactions on Systems, Man, and Cybernetics Part C Applications and Reviews*, *31*(2), 216-233.

Balbudhe, K. S., Kulkarni, N., Waghole, S., Datar, A., & Khanna, K. (2016). A Survey on Handwritten Character. *International Journal of Innovative Research in Computer and Communication Engineering, 4*(2), 2322-2327.

Bengio, Y. (2009). Learning deep architectures for AI. *Foundations and Trends® in Machine Learning*, *2*(1), 1-127.

Berkeley-Linguistics. (2018). *Script Encoding Initiative*. Retrieved June 4, 2018, from http://linguistics.berkeley.edu/sei/progress-overview.html

Besekar, D. N. (2012). Special approach for recognition of handwritten MODI script's vowels. *International Journal of Computer Applications, (IJCA), MAHA, MEDHA*(1), 48-52.

Besekar, D. N. (2013). Study for theoretical analysis of handwritten MODI script - a recognition perspective. *International Journal of Computer Applications*, *64r*(3), 45-49.

Besekar, D. N., & Ramteke, R. J. (2012). Feature extraction algorithm for handwritten numerals recognition of MODI script using zoning-based approach. *International Journal of Systems, Algorithms & Applications, 2*(November), 2-5.

Bhattacharya, U., Shridhar, M., & Parui, S. K. (2006). On Recognition of Handwritten Bangla Characters. In *Proceedings of the 5th Indian Conference on Computer Vision, Graphics and Image Processing* (pp. 817-828). Heidelberg, Germany: Springer-Verlag.

Capin, T., & Pulli, K. (1990). The State of the Art in online handwriting recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *12* (August), 787-808.

CDAC. (2016). *Modi Script Tools*. Retrieved June 6, 2018, from https://www.cdac.in/index.aspx?id=pk_itn_spot1006

Chacko, B. P., Krishnan, V. V, Raju, G., & Anto, B. P. (2012). Handwritten character recognition using wavelet energy and extreme learning machine. *International Journal of Machine Learning and Cybernetics*, *3*(2), 149-161.

Chandure, S. L., & Inamdar, V. (2016, December 19-21). Performance analysis of handwritten devnagari and MODI character recognition system. In *International Conference on Computing, Analytics and Security Trends (CAST)* (pp. 513-516). Pune, India.

Chaudhuri, A., Mandaviya, K., Badelia, P., & Ghosh, S. K. (2017). Optical character recognition systems. In *Optical Character Recognition Systems for Different Languages with Soft Computing* (pp. 9-41). Cham, Switzerland: Springer.

Chen, Y. K., & Wang, J. F. (2000). Segmentation of single- or multiple-touching handwritten numeral string using background and foreground analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *22*(11), 1304-1317.

Chherawala, Y., Roy, P. P., & Cheriet, M. (2016). Feature set evaluation for offline handwriting recognition systems : application to the recurrent neural network model. *IEE Transactions on Cybernetics*, *46*(12), 2825-2836.

Cires, D. C., Meier, U., & Gambardella, L. M. (2011). Convolutional neural network committees for handwritten character classification. *International Conference on Document Analysis and Recognition*, *10*, 1135-1139.

Deng, L. (2014). Deep learning: methods and applications. *Foundations and Trends in Signal Processing*, *7*(3-4), 197-387.

Deshmukh, M., Patil, M., & Kohle, S. (2017). The divide-and-Conquer based algorithm to detect and correct the skew angle in the old age historical handwritten MODI lipi documents. *International Journal of Computer Science and Applications*, *14*(2), 47-63.

Deshpande, P. (2013, February 13). Rare manuscripts to be digitised by archives dept. *Mid Day India News*. Pune. Retrieved June 6, 2018, from http://archive.mid-day.com/news/2013/feb/ 130213-pune-rare-manuscripts-to-be-digitalised-by-archives-dept.htm.

Dhaka, V. P., & Sharma, M. K. (2015). An efficient segmentation technique for Devanagari offline handwritten scripts using the Feedforward Neural Network. *Neural Computing and Applications*, *26*(8), 1881-1893.

Garain, U., & Chaudhur, B. B. (2001, September 13). Segmentation of touching characters in printed Devnagari and Bangla scripts using fuzzy multifactorial analysis. In *Proceedings of Sixth International Conference on Document Analysis and Recognition* (pp. 805-809). Seattle, WA, USA.

Gharde, S. S., & Ramteke, R. J. (2016, December 22-24). Recognition of characters in Indian MODI script. In *International Conference on Global Trends in Signal Processing, Information Computing and Communication (ICGTSPICC)* (pp. 236-240). Jalgaon, India.

Goswami, R., & Sharma, O. P. (2013). A review on character recognition techniques. *International Journal of Computer Applications*, *83*(7), 19-23.

Hindu. (2007). House library panel visits Saraswathi Mahal. *The Hindu*. Tanjavur. Retrieved June 4, 2018, from thehindu.com/todays-paper/tp-national/tp-tamilnadu/house-library-panel-visits-saraswathimahal / article1924551.ece

Hindu. (2013). MODI documents of Maratta era to be digitised. *The Hindu*. Tanjavur. Retrieved June 4, 2018, from thehindu.com/todays-paper/tp-national/MODI-documents-of-maratta-era-to-be-digitised/ article4738336.ece

Hinton, G. (2007). Learning multiple layers of representation. *Trends in Cognitive Science*, *11*(10), 428-434.

Hinton, G., & Salakhutdinov, R. (2006). Reducing dimensionality of data with neural network. *Science*, *313*(5786), 504-507.

Hinton, G., Deng, L., Yu, D., Dahl, G., Mohamed, A., Jaitly, N., … & Kingsbury, B. (2012). Deep neural networks for acoustic modeling in speech recognition. *IEEE Signal Processing Magazine*, *29*, 82-97.

Jain, A. K., Duin, R. P. W., & Mao, J. (2000). Statistical pattern recognition: a review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *22*(1), 4-37.

Jayadevan, R., Kolhe, S. R., Patil, P. M., & Pal, U. (2011). Offline recognition of devanagari script: A survey. *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews*, *41*(6), 782-796.

Jindal, M. K., Sharma, R. K., & Lehal, G. S. (2009, January 09-10). Segmentation of touching characters in upper zone in printed Gurmukhi script. In *Proceedings of the 2nd Bangalore Annual Compute Conference* (p. 9). Bangalore, India.

John, J., Pramod, K. V., & Balakrishnan, K. (2011, February 19-20). Handwritten character recognition of South Indian scripts : a review. In *Proceedings of National Conference on Indian Language Computing* (pp. 1-6). Kochi, India.

Jyoti, & Kumar, A. (2016). Thresholding Techniques for Ancient Document Images. *International Journal of Advanced Research in Computer Engineering & Technology*, *5*(9), 2345-2350.

Kale, K. V., Patil, P. D. D., Chavan, S. V., Kazi, M. M., & Rode, Y. S. (2014). Zernike moment feature extraction for handwritten devanagari compound character recognition. *International Journal of Advanced Research in Artificial Intelligence*, *3*(1), 459-466.

Khillari, R. (2008). *History of MODI script*. Retrieved June 6, 2018, from http://modi-script.blogspot.com/2008/05/history-of-modi-script.html

Kim, I. J., & Xie, X. (2015). Handwritten hangul recognition using deep convolutional neural networks. *International Journal on Document Analysis and Recognition (IJDAR)*, *18*(1), 1-13.

Kulkarni, S., Borde, P., Manza, R., & Yannawar, P. (2014). Recognition of handwritten MODI numerals using Hu and Zernike features. *Computer Vision and Pattern Recognition, 1*, 1-7.

Kulkarni, S., Borde, P., Manza, R., & Yannawar, P. (2015a). Review on recent advances in automatic handwritten MODI script recognition. *International Journal of Computer Applications*, *115*(19), 975-8887.

Kulkarni, S., Borde, P., Manza, R., & Yannawar, P. (2015b). Analysis of orthogonal moments for recognition of handwritten modi numerals. *VNSGU Journal of Science and Technology, 4*(1), 36-43.

Kulkarni, S., Borde, P., Manza, R., & Yannawar, P. (2016). *Recognition of Partial Handwritten MODI characters using zoning*. Paper presented at National Conference of Image Processing and Pattern Recognition, Bidar, Karnataka. (Unpublished)

Kulkarni, S., Borde, P., Manza, R., & Yannawar, P. (2017, May 13). Text Line Segmentation of Handwritten Historical MODI Documents. In *International Conference on Information, Communication and Computing Technology (ICICCT2017)* (pp. 1–12). New Delhi, India.

Larochelle, H., Courville, A., & Bergstra, J. (2007, June 20-24). An empirical evaluation of deep architectures on problems with many factors of variation. *Proceedings of International Conference on Machine Learning* (pp. 473-480). Corvalis, Oregon, USA.

Liu, C. L., & Fujisawa, H. (2008). Classification and learning methods for character recognition: Advances and remaining problems. In *Machine Learning in Document Analysis and Recognition* (pp. 139-161). Heidelberg, Germany: Springer.

Mahnaz, S., & Maher, S. A. (2015). Skew detection and correction based on an axes-parallel bounding box. *International Journal on Document Analysis and Recognition (IJDAR)*, *18*(1), 59-71.

Maitra, D. Sen, Bhattacharya, U., & Parui, S. K. (2015, August). CNN based common approach to handwritten character recognition of multiple scripts. In *2015 13th International Conference on Document Analysis and Recognition (ICDAR)* (pp. 1021-1025). Tunis, Tunisia.

Maurya, R. K., & Maurya, S. R. (2018). Recognition of a medieval indic-MODI script using empirically determined heuristics in hybrid feature space. *International Journal of Computer Sciences and Engineering, 6*(2), 136-142.

Oujaoura, M., El Ayachi, R., Fakir, M., Bouikhalene, B., & Minaoui, B. (2012). Zernike moments and neural networks for recognition of isolated Arabic characters. *International Journal of Computer Engineering Science*, *2*(3), 17-25.

Pandey, A. (2011). *Final Proposal to Encode the Khojki Script in ISO / IEC 10646*. Retrieved June 6, 2018, from https://www.unicode.org/L2/L2011/11212r2-n4034-MODI .pdf

Patil, P. A. (2016). Character recognition system for MODI script. *International Journal of Computational Engineering Research (IJCER), 6*(9), 31-38.

Ramraje, R. A. (2013). History of MODI script in Maharashtra. *Global Online Electronic International Interdisciplinary Research Journal (GOEIIRJ)*, *2*(1), 54-58.

Ramteke, A., & Katkar, G. (2013). Recognition of off-line MODI Script : a structure similarity approach. *Proceedings International Journal of ICT and Management*, *1*(1), 12-15.

Ramteke, A., & Rane, M. (2012). A survey on offline recognition of handwritten devanagari script. *International Journal of Scientific and Engineering Research*, *3*(5), 782-796.

Rathi, S. (2015). Recognition and conversion of handwritten MODI characters. *International Journal of Technical Research and Applications*, *111*(2), 28-34.

Rehman, A., & Saba, T. (2011). Performance analysis of character segmentation approach for cursive script on benchmark database. *Digital Signal Processing*, *21*(3), 486-490.

Sadanand, K., Borde, P. L., Ramesh, M., & Pravin, Y. (2015a). Offline MODI character recognition using complex moments. *Procedia Computer Science*, *58*, 516-523.

Sadanand, K., Borde, P., Ramesh, M., & Pravin, Y. (2015b, September 10-12). Impact of zoning on zernike moments for handwritten MODI character recognition. In *2015 International Conference on Computer Communication and Control (IC4)* (pp. 1-6). Indore, India.

Savant, M., & Rakshana, K. (2005). Marathi Modi Script: Origin, Evolution and Significance. *A Bi-Monthly Publication of the National Mission for Manuscript*, *1*(1), 16-28.

Sepahvand, M., Abdali-Mohammadi, F., & Mardukhi, F. (2017). Evolutionary metric-learning-based recognition algorithm for online isolated persian / arabic characters, reconstructed using inertial pen signals. *IEE Transactions on Cybernetics*, *47*(9), 2872-2884.

Sharma, P. (2013). Classification in pattern recognition : a review. *International Journal of Computational Engineering Research (IJCER)*, *3*(4), 298-306.

Shelke, S. S., & Apte, S. S. (2011). A multistage handwritten marathi compound character recognition scheme using neural networks and wavelet features. *International Journal of Signal Processing, Image Processing and Pattern Recognition*, *4*(1), 81-94.

Tamhankar, P. A., & Kolhe, S. R. (2018). Skew detection, skew correction, and shirorekha extraction of line segmented offline handwritten historical marathi documents written in MODI. *International Journal of Latest Trends in Engineering and Technology, 10*(1), 344-348.

Thakre, R. (2016). *Reviving MODI script*. Retrieved June 4, 2018, from http://shivajiscript.blogspot.com/

Times. (2014). Band of researchers, enthusiasts strive to keep MODI script alive. *Times of India*. Pune. Retrieved June 6, 2018, from www.timesofindia.indiatimes.com/city/ pune/Band-of-researchers-enthusiasts-strive-to-keep-MODIscriptalive/ articleshow/30761335.cms

Tripathy, J. (2010). Reconstruction of oriya alphabets using zernike moments. *International Journal of Computer Applications*, *8*(8), 26-32.

Zhang, X. Y., Yin, F., Zhang, Y. M., Liu, C. L., & Bengio, Y. (2018). Drawing and recognizing chinese characters with recurrent neural network. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *40*(4), 849-862.

Pertanika J. Sci. & Technol. 27 (4): 1671 - 1692 (2019)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Resolution Adaptive Threshold Selection for Gradient Edge Predictor in Lossless Biomedical Image Compression

Urvashi[1]*, Meenakshi Sood[2] and Emjee Puthooran[3]

Department of Electronics and Communication Engineering, Jaypee University of Information Technology, Waknaghat, India

ABSTRACT

The high-resolution digital images generated for medical diagnosis produce the extremely large volume of digital data. This necessitates the use of image compression for medical data to be processed, archived and transmitted through a computer network in an efficient way. Due to the criticality in disease diagnostics and legal reasons, biomedical images require lossless compression to prevent permanent loss of image data. Among various approaches to lossless compression of medical images, predictive coding techniques have high coding efficiency and low complexity. Gradient Edge Predictor (GED) used in predictive coding technique for prediction has higher coding efficiency as compared to Median Edge Detector (MED) used in JPEG-LS. GED has lower computational complexity as compared to Gradient Adaptive Predictor (GAP) used in CALIC. GED is a threshold based predictor, however there is no specific method adopted in literature to decide the threshold value for prediction. This paper presents an efficient prediction solution based on predictive coding technique. The main objective of this research work is to develop a Resolution Independent Gradient Edge Predictor (RIGED) technique for choosing an optimal threshold value for GED predictor which will give minimum entropy value irrespective of the type of modality and resolution of the medical images. The empirical experimentation and analysis gave percentage improvement

ARTICLE INFO

Article history:
Received: 3 July 2018
Accepted: 23 April 2019
Published: 21 October 2019

E-mail addresses:
survashi2793@gmail.com (Urvashi)
meenakshi.sood@juit.ac.in (Meenakshi Sood)
emjee.puthooran@juit.ac.in (Emjee Puthooran)
**Corresponding author*

ISSN: 0128-7680
e-ISSN: 2231-8526

of the proposed model as 32.4% over MED and percentage difference between high complexity GAP and proposed predictor as 0.68 % in terms of entropy for medical image dataset of different modalities having different resolution.

Keywords: Lossless image compression, medical imaging, predictive coding, predictor, resolution independent gradient edge detection

---

INTRODUCTION

With the advancement of information and communication technology, the amount of digital data produced all over the world is tremendous. This is especially true for biomedical images as typical terabytes of digital image data is generated by a hospital every year (Placidi, 2009). Medical imaging is experiencing a growth in terms of usage and image resolution, specifically in diagnostics systems that entail a large volume of medical images. The higher resolution of medical images results in large amount of digital data which may need to be stored, processed and transmitted in an efficient way (Khatkar & Kumar, 2018). These facts make a demand for effective techniques of image compression. Efficient compression techniques help to reduce bandwidth and storage space requirement. Image compression is a very important aspect of efficient transmission and archiving of medical image data along with the successful real-time application of telemedicine.

Many hospitals have small clinics and satellite centres located in the desolate areas where distance is a critical issue to deliver the health care services. Patient inhabiting in desolate, remote and semi-urban areas find hard time to travelling to the hospitals especially for diagnostic purpose. For the convenience of patients suffering from severe diseases, the big hospitals make use of telemedicine and tele-radiology applications providing health care facilitates from distance in such areas. It allows the physician at the remote centers to take a medical image data of patients (MRI or CT scan) and transmit it to the radiologist in the main hospital placed in the city. Radiologist examines the medical image data of patients and sends back the diagnostic details to the physician. In addition to telemedicine application in medical field, compression techniques are useful in remote sensing applications.

Removal of redundancies present in a medical image results in reduced number of bits to represent the information and achieves compression. The efficiency of the image transmission system is improved by compression techniques that can reduce data size and transmission time. General techniques used for image compression are lossy and lossless. Lossy compression is not generally acceptable to be used in medical field although it results in higher compression ratio, because a permanent loss of biomedical data during compression cannot be tolerated. However, lossless compression algorithms can produce images which are an exact replica of the original images without any loss of diagnostic information. Lossless compression of the image is important and appropriate for medical images because any loss of data could affect the clinical diagnostic process and it may lead to wrong diagnosis (Kaur et al., 2015). Many compression techniques are available in literature viz., transformation coding, entropy encoding and dictionary encoding (Sharma et al., 2017). Among the different techniques, predictive coding performs well for lossless coding in terms of coding efficiency and compression time. The efficiency of predictive coding depends upon the predictors (Avramovicl & Savicl, 2011).

RELATED WORK

Many researchers have worked on applying predictive coding technique for lossless compression of medical images.

2D predictors used in predictive coding technique for reducing interpixel redundancy from the image and is operated in frame by frame basis. JPEG-LS (Joint Photographic Expert Group-Lossless Compression) (ISO/IEC, 1999) and CALIC (Context-Based Adaptive Lossless Image Coding) are 2D benchmark algorithms for prediction based compression. JPEG-LS mainly consists of context modeling, pixel prediction, and prediction error encoding and uses Median Edge Detection (MED) predictor (Avramovic, 2012). CALIC employs Gradient Adaptive Predictor (GAP). The well-known MED predictor is simple but has less coding efficiency than computationally complex GAP (Wu & Memon, 1996). Avramovicl and Savicl (2011) proposed prediction based algorithms on the estimation of local gradients and detection of edges. Standard predictors MED and GAP were analyzed and entropies of predicted medical images were also compared. Baware and Save (2016) presented GAP predictor of prediction scheme that was adaptive to gradients defined in four directions. Errors obtained were grouped on the basis of max plane coding before entropy encoding, which enhanced coding efficiency. This scheme is compared with CALIC and DPCM (Differential Pulse Code Modulation) methods achieve better results in terms of compression ratio, bit rates and lesser computational complexity (Baware & Save, 2016). Avramovic and Reljin (2010) proposed a threshold controlled gradient edge detection (GED) for lossless image compression which combined simplicity of MED predictor and efficiency of GAP. GED predictor achieving bit rates comparable to more complex GAP predictor. It is shown that on the selected set of original medical images, GED predictor shows better results. Shrikhande and Bairagi (2013) introduced many techniques for achieving data compression. It was concluded by author that CALIC algorithm that used GAP predictor achieved better results in terms of entropy values but under higher time (and space) complexity as compared to JPEG-LS that employed MED predictor. Al-Mahmood and Al-Rubaye (2014) presented a lossless image compression technique for compressing medical and natural images by using the combination of adaptive predictive coding and bit plane slicing. In this technique most significant bits used adaptive predictive coding while other used run length coding. This technique achieved higher compression ratio for lossless coding that guaranted full recovery of an image. Fouad proposed a lossless image compression technique integrating an integer wavelet transform with a prediction step (Fouad, 2015). It was a simple median edge detector algorithm used with JPEG-LS technique. Image was first transformed using the predictor and an error image was obtained. The error image was then entropy encoded after being transferred through an integer wavelet transform. Higher compression ratio was achieved by this technique than competing techniques.

Hashemi-Berenjabad and Mojarrad (2016) presented a brief review of medical image compression methods which were applicable in telemedicine, e-health, and teleconsultation. Researchers tried to improve compression techniques in the case of compression ratio and quality of reconstruction images system. Mofreh and Refaat (2016) proposed a new hybrid lossless image compression that depended upon predictive coding, DWT transformation and Huffman coding. Output of predictor is fed to DWT transformation and then Huffman encoding is applied to output of DWT. This hybrid scheme gave better results in terms of bit rate as compare to Huffman and Huffman + DWT. Thombare1 proposed a technique that used a low complexity, lossless, compression algorithm for the compression of 3-D volumetric medical images. 3-dimensional nature of the data was exploited by using 3-D linear prediction (Thombare, 2016). Shanmathi and Maniyath (2017) analyzed commonly used predictors; MED, GAP and GED in term of entropies after prediction. Output of prediction error image was measured for the efficiency of predictor. Analysis showed that GED provided better results as it was simple to implement and also provided efficient results. Context modeling and entropy encoding was also explained by the author. Prediction error was further compressed in lossless manner by using arithmetic coding technique. The lossless compression technique achieved higher PSNR and lower CR values as compared to near lossless image compression method.

Research Gaps

In literature different predictors are described for predictive coding and it is concluded that GED gives better results in terms of entropy at low computational complexity. GED is based on threshold value and for efficient prediction; threshold value should be chosen to provide minimum entropy of residual image. However, in literature, no threshold value is specified for prediction and existing predictor is resolution and modality dependent. Threshold value is different for varying image resolutions and choice of threshold is difficult as there is no specific method reported earlier. But, it is necessary to find an optimal threshold value that provides minimum entropy for varying image modalities and resolutions. The proposed RIGED is designed to provide an optimal threshold value independent of image resolution and modality.

Optimal threshold value is a universal value that does not depend upon the image resolution and modality for prediction in predictive coding technique and also provides minimum entropy value irrespective of image resolution and modality. Dataset used by different authors in literature is small for different modalities of medical images.

Objectives

The main objective of this research work was to develop a technique for choosing an efficient threshold value for GED predictor irrespective of the type of modality adopted and resolution of the medical image. An exhaustive empirical work had been done on various large datasets of different resolutions and of different modalities containing the varying number of slices. Resolution Independent Gradient Edge Predictor (RIGED) was developed in this paper that provided efficient prediction in terms of entropy at optimal threshold for 8-bit depth medical images having varying resolution and modalities.

Entropy symbolizes the number of bits required to encode for compression so the whole of research work is based on a measure of entropy. The novelty of this research paper is to come up with an optimal universal threshold value for prediction and to get the residual image with the lowest entropy. Minor entropy variations at the optimal threshold value for different image modalities and resolutions are taken care of. The optimal threshold value is independent of modality as well as image resolution for every 8-bit image depth.

MATERIALS AND METHODS

Data Set

Medical images datasets were collected from Center for Image Processing Research (CIPR), Massachusetts General Hospital (MGH) , Micro-Dicom , OsiriX sources. All the images were 8-bit depth. They were used for examining predictor efficiency testing and validation of different predictors.

Predictive Based Lossless Image Compression

The amount of information present in an image that does not provide any relevant information is called redundancy (Avramovic, 2012). Spatial and temporal redundancy present in an image should be removed from the image for better compression. Correlation between neighboring pixels in image results in interpixel spatial redundancy. Difference between adjacent pixels can be used to represent an image to reduce the interpixel redundancy. Steps of data processing contained by predictive based lossless image compression are a prediction, context modeling and entropy encoding as shown in Figure1 (Me et al., 2012).

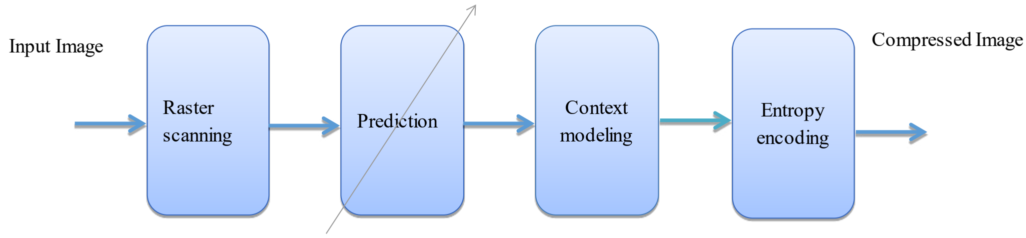


Figure1. General scheme of lossless predictive image compression

Prediction is the essential part of the image compression as it removes most of the redundancy from the image. Prediction of every pixel individually can be done from a context (a group of neighboring pixels) in a raster scan order. Residual image is obtained by subtracting predicted image from an original image. The residual image has lower entropy as compared to the original image. Therefore lesser amount of bits is required to encode it. For the efficiency of compression methods, choice of an optimal predictor is important. Predictor's efficiency is based on the entropy of the prediction error. Lower the entropy better will be the performance of the predictor.

2-D predictors. Spatial redundancy in 2D images can be removed by 2D predictors that exploit the redundancy from the image. 2D predictors can also be used for 3D image compression operating in a frame by frame basis. Here 3D volumetric data was split into 2D images and encoded individually. Common Scheme for labeling of causal neighbors in 2D predictors was shown in equation [1]. Let Xi,j is a current pixel for prediction and XN (North), XW (West), XWW (West-West), XNW (North-West), XNE (North-East), XNN (North-North) and XNNE (North-North-East) are neighboring pixels of Xi, j. Neighboring pixels are denoted as follows:

$$\begin{aligned}&XN = X[i, j-1],\ XW = X[i-1, j],\\ &XNW = X[i-1, j-1],\ XNE = X[i+1, j-1]\\ &XNN = X[i, j-2],\ XNW = X[i-2, j],\\ &XNNE = X[i+1, j-2]\end{aligned} \quad [1]$$

The various 2D predictors employed are Median Edge Predictor (MED), Gradient Adjusted Predictor (GAP) and Gradient Edge Predictor (GED) which are used in this research work are briefed in the following section.

---

1. "CIPR." [Online].Available: http://www.cipr.rpi.edu/resource/sequences/sequence01.html.
2. "MGH)." [Online]. Available: http://www.cma.mgh.harvard.edu.ibsr.
3. "Micro-Dicom Dataset." [Online]. Available: http://www.microdicom.com/downloads.html.
4. "OsiriX Dataset." " [Online]. Available: http://pubimage.hcuge.ch:8080/.

Median Edge Detection Predictors (MED). MED predictor was used in JPEG-LS. Three causal pixels were used to select one of the three sub-predictors depending upon whether it was vertical edge or horizontal edge (Weinberger et al., 2000). MED selected the median value among neighboring pixels N, W and W+N-NW.

Gradient Adjusted Predictors (GAP). GAP is based on the gradient estimation which can adapt itself to the intensity gradients of immediate neighbors of predicted pixel. It estimates three types of edges, simple, sharp and a weak edge. To estimate the local gradients and determine the prediction values on some threshold, six causal pixels were used. (Heuristic values of threshold are 8, 32 and 80) (Weinberger et al., 2000) Common scheme for labeling pixels is shown in equation [1].

Gradient Edge Detection (GED). GED predictor is best combination of simplicity and efficiency. Like GAP, it uses local gradient estimation on proper threshold (T) value and chooses between three sub predictors, defined as in MED predictor (Me et al., 2012). It selects one threshold and it can be predefined or user-defined (Avramovicl & Savicl, 2011).

$$
\begin{aligned}
&Av = |NW-W| + |NN-N| \\
&Ah=|WW-W|+|NW-N| \\
&\text{if } Av-Ah > T, \qquad Px = W \\
&\text{else if } Av-Ah < -T, \qquad Px = N \\
&\text{else } Px = N+W-NW \\
&\text{where, } T = \text{Threshold value}
\end{aligned} \qquad [2]
$$

Proposed Model

Resolution Independent Gradient Edge Detection [RIGED]. Existing GED predictor is the finest combination of simplicity and efficiency. It takes advantage from state-of-the-art MED and GAP predictors. However, it is a threshold based predictor and there is no specific method in literature for its threshold value selection.

Threshold selection is important for efficient prediction and efficiency can be increase by reducing the entropy of residual image. Residual image is further entropy encoded for removing coding redundancy. Lower entropy residual requires lower number of bits to encode it. Efficiency of predictor is improved by making it resolution and modality independent. RIGED is the extension of existing threshold based GED predictor which is subjective in nature. Choosing the value of threshold for image prediction is not reported in literature. RIGED removes its demerit by providing an optimal threshold value.

Prediction done at optimal threshold value provides efficient results in terms of entropy. RIGED is less computationally complex and provides better results as that of highly efficient GAP in terms of entropy. RIGED was tested for different image resolution and modalities collected from different standard medical image dataset sources. From the extensive experiments conducted on the medical image datasets of varying modalities and varying resolution, optimal threshold value was selected empirically.

Residual image resulting from RIGED has minimum entropy value resulting in lesser number of bits required for encoding. With proper selection of threshold value for prediction, RIGED is less computationally complex and gives comparable results as that of highly efficient GAP predictor.

$$\left.\begin{array}{c} T = 32\ if\ dataset\ resolution = 256 \times 256 \\ T = 64\ if\ dataset\ resolution = 512 \times 512 \\ T = 44\ \ for\ set\ of\ different\ resolution \\ images\ between\ 256 \times 256\ and\ 512 \times 512 \\ i.e., 256 \times 256 < T < 512 \times 512 \end{array}\right\} \quad [3]$$

where, T = Threshold value

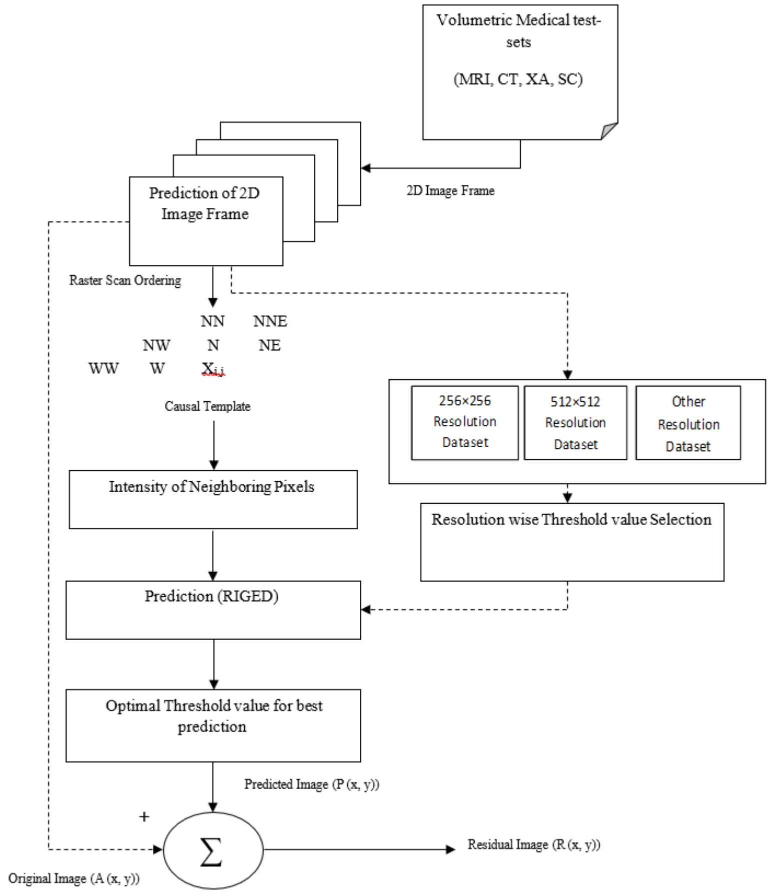


Figure 2. Resolution independent gradient edge detection

A 2D image frame was taken from the volumetric image data set containing number of image frames and the processing of image slices was done on frame-by-frame basis. Neighboring pixels were extracted using causal template of the image and the pixel prediction was done in raster scan order. Pixel prediction was done by RIGED and a common threshold value was used that was optimal for 8 bit depth images. After prediction, residual image or a prediction error image R (x, y) was obtained by subtracting predicted image P (x, y) from original image A (x, y). R (x, y) obtained after predicting at optimal threshold value provided lower entropy that was further entropy encoded. Lower entropy of R (x, y) requires fewer numbers of bits to encode it and provides high compression efficiency. Entire process of the proposed algorithm was repeated for each frame resulting in efficient prediction measured in terms of entropy.

The empirical experimentations involved the number of experiments detailed in this section.

Experiment 1. In this experiment the efficiency of predictor was calculated for medical images of different modalities having resolution 256×256 taken from CIPR and MGH sources. Aim of this experiment was to find the optimal threshold value for prediction of standard size medical images having resolution 256×256. Prediction was done slice by slice basis from 2D predictor and average entropy of residual image was calculated for each modality of 256×256 resolution. The dataset consisted of the following images as shown in Table 1.

Table 1

Dataset detail of R1 (Resolution 256×256)

TAG	Sequence Name	Modality	Image Size	Slices
CIPR-CT-01	CT_Aperts	CT	256×256	97
CIPR-CT-02	CT_carotid	CT	256×256	74
CIPR-CT-03	CT_skull	CT	256×256	203
CIPR-CT-04	CT_wrist	CT	256×256	183
CIPR-MR-01	MR_liver_t1	MR	256×256	77
CIPR-MR-02	MR_liver_t2e1	MR	256×256	58
CIPR-MR-03	MR_ped_chest	MR	256×256	58
CIPR-MR-04	MR_sag_head	MR	256×256	58
MGH-MR-01	657_10	MR (PD)	256×256	18
MGH-MR-02	657_2	MR (T1)	256×256	18
MGH-MR-03	657_11	MR (T2)	256×256	18

Experiment 2. Next set up of experiment involved calculation of entropy with different thresholds of predictor. As image resolutions of 256×256 and 512×512 were standard size for medical images so the objective of this experiment was to find the optimal threshold value that provided the minimum entropy of residual image after prediction. They were tested on various images used in this set of experiment includes MR, CT and XA images. These different image modalities having resolution 512×512 were taken from MIDI and OSRX sources. The dataset consisted of the following images as shown in Table 2.

Table 2

Dataset detail of R2 (Resolution 512×512)

TAG	Sequence Name	Modality	Image Size	Slices
MIDI-MR-01	SAG-T1	MR	512×512	13
MIDI-MR-02	COR-T1	MR	512×512	20
MIDI-MR-03	COR-FSE-T2	MR	512×512	20
MIDI-MR-04	COR-FLAIR	MR	512×512	12
MIDI-MR-05	AX-FSE-T2	MR	512×512	18
MIDI-MR-06	COR-T1-POST-GAD	MR	512×512	15
MIDI-MR-07	SAG-T1-POST-GAD	MR	512×512	15
OSRX_CT_01	BREBIX	CT	512×512	245
OSRX_CT_01	MAGIX	CT	512×512	77
OSRX_MR_01	BRANIX	MR	512×512	22
OSRX_XA__11	GUSERAMBIX	XA	512×512	14
OSRX_XA_01	GUSERAMBIX	XA	512×512	76
PHNT_MR	E1154S7I	MR	512×512	76

Experiment 3. Dataset of medical images having resolution other than 256×256 and 512×512 were also taken from OSRX and tested by predictor with varying values of threshold. To obtain the optimal threshold value for the set of medical images of different resolutions other than 256×256 and 512×512 was the main objective of this experiment. A common threshold value was selected that was optimal for every possible combination of resolution. The dataset consisted images of different resolutions other than 256×256 and 512×512 as shown in Table 3.

Table 3

Dataset detail of R3 (Resolution other than 256×256 and 512×512)

TAG	Sequence Name	Modality	Image Size	Slices
OSRX_CT_02	CEREBRIX	CT	336×336	83
OSRX_SC_01	CEREBRIX	SC	270×320	28
OSRX_SC_02	CEREBRIX	SC	270×320	24
OSRX_SC_04	CEREBRIX	SC	320×384	29

Experiment 4. In this experiment same set of medical images including MR, CT, XA and SC having different resolution and different field of view were tested by state-of-the art MED and GAP predictors. Objective of this experiment was to verify the improvement of proposed RIGED over MED and GAP. RIGED was compared with less computationally complex MED and highly efficient GAP predictor in terms of entropy, computational complexity and execution time.

Performance Metrics

Different predictors for medical images were examined with performance metrics as entropy and computational complexity. Entropy is the main parameter to evaluate the predictor for lossless predictive coding.

Entropies of prediction. Number of bits used to represent the information of an image is described by entropy (Al-Naqeeb & Nordin, 2017). After removing the spatial redundancy from the image by using predictors, entropy is calculated that can be used for the estimation of a final compression ratio (Puthooran et al., 2013). Let X is the random variable of an image, with an alphabet Y= (y_0, y_1…..y_{n-1}) which means it is N-bit image. Entropy of an image is calculated as:

$$H(X) = -\sum_{x \epsilon Y} p(x) \log p(x) \qquad [4]$$

Where, p(x) is probability of a symbol X.

Computational Complexity. A number of operations required for implementation of predictor algorithm compute the computational complexity and time period required for executing the program. Higher the complexity of predictor more will be the running time. A predictor is highly efficient if it provides good results with minimum complexity and runs time.

RESULTS AND DISCUSSION

Comparative Analysis of Different Predictors

In the beginning, a comparative analysis was done for the different techniques. Results of medical images from different predictors MED, GAP and GED as obtained are shown in Figure 3 and Figure 4. It was found that interpixel redundancy was better removed by GAP than MED predictor in terms of entropy of prediction error image. Results of GED predictor were approximately same as that of GAP predictor though it was computationally simpler than GAP predictor. Original image and predicted images that were obtained from MED, GAP and GED predictors with their histograms are depicted in Figure 3 and histograms of residual images are shown in Figure 4.

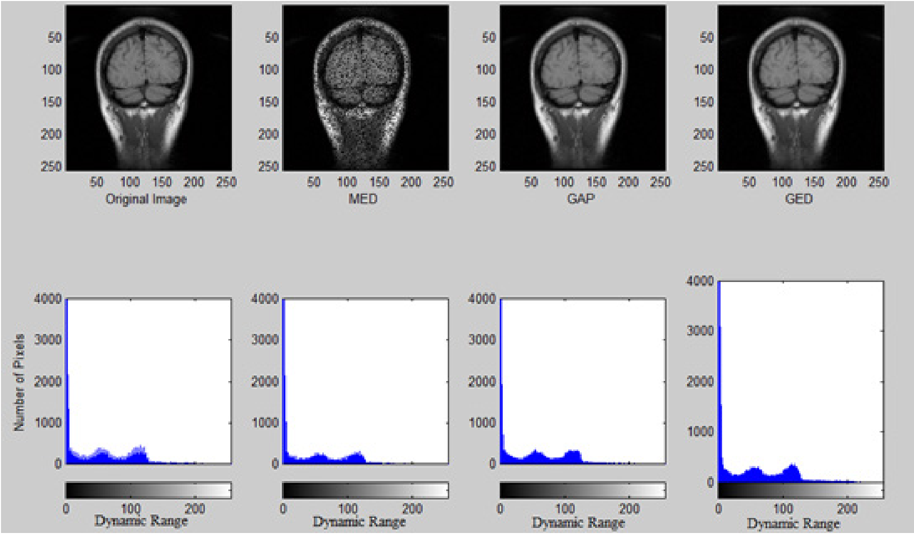


Figure 3. Original MRI-scan image of brain, predicted image obtained from MED, GAP and GED predictors and their histograms

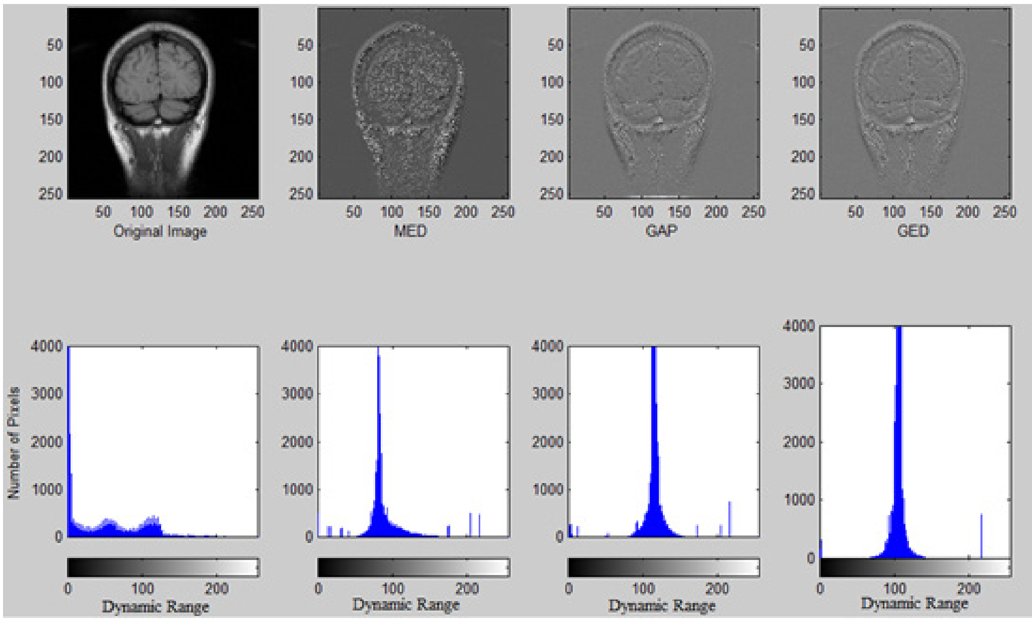


Figure 4. Original MRI-scan image of brain, residual image obtained from MED, GAP and GED predictors and their histograms

Analysis of Entropy Values for Different Resolutions

Structural analysis was carried out for various samples of medical images from all predictors MED, GAP and GED, which provided better assessment with respect to entropy and computational complexity. GED predictor was based upon one threshold value which was selected manually for better results. For images with different resolution, different threshold value had to be chosen for optimal results. Our proposed algorithm RIGED was designed to make this algorithm independent of the modality and resolution. RIGED was tested for different thresholds from 8- 256. After calculating the entropy for a threshold 2k, it was observed that there was significant change in entropy when threshold changed from 8 to 32 whereas threshold ranging from 32 to 80 there was minor variation in the value of entropy. RIGED prediction was also calculated for threshold values 32 to 128 with a difference of 16 (48, 64, 80, 96 and 112). Nomenclature used in this paper was RIGED_8 representing GED predictor with threshold value = 8.

Entropy for R1 dataset for varying threshold values. All the images from different modalities had varying number of slices. Weighted average of entropy for all the slices was calculated. It was found that for set of medical image samples having resolution 256×256 the lowest entropy was obtained at threshold value of 32 i.e., optimum value of threshold was 32 for images of 256×256 resolution. Graphical representation of entropy obtained for R1 set (256×256 resolution) of medical image dataset from GED predictor at varying threshold value is shown in Figure 5. For R1 set of medical images, highest entropy was obtained at the threshold value of 8. On increasing value of threshold, entropy decreased till threshold value of 32 and it started increasing thereafter. There was slight variation in entropy value for threshold values of 32 to 48 and after that significant change was observed for higher values of threshold after 48 as shown in Figure 5.

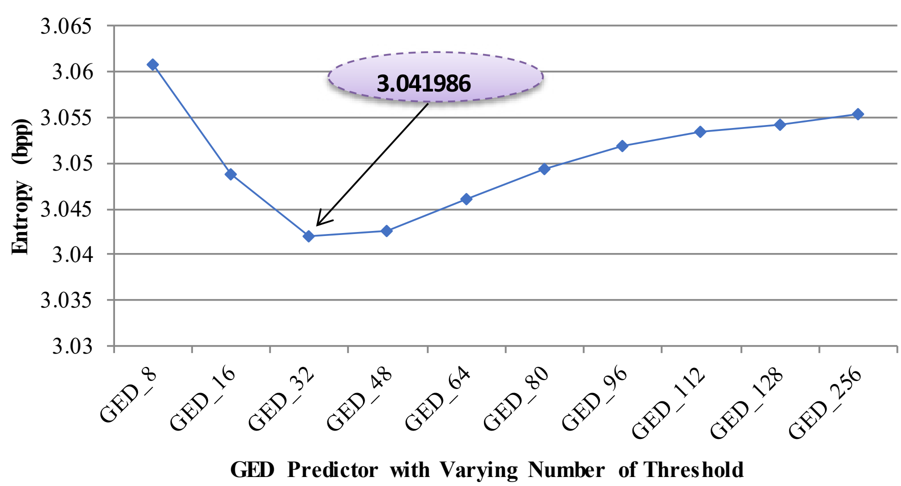


Figure 5. Average value of entropy from GED predictor with varying threshold values for R1 resolution.

Entropy for R2 Dataset for Varying Threshold Values. As stated in experiment 2, the GED predictor was used to determine the optimal threshold value for a set of medical image samples having resolution 512×512. Set of R2 (resolution 512×512) contained different modalities of medical images with number of slices. Entropy was calculated for all slices for different modalities of images and weighted average of entropy for R2 set was obtained. As the resolution increased from 256×256 to 512×512, entropy decreased at varying value of threshold from 32 to 64 and then increased for higher values of threshold of 64, so it was found that minimum value of entropy was obtained at threshold value of 64 for the group R2 of medical image dataset. The results are depicted in Figure 6 showing the minimum entropy at threshold value of 64.

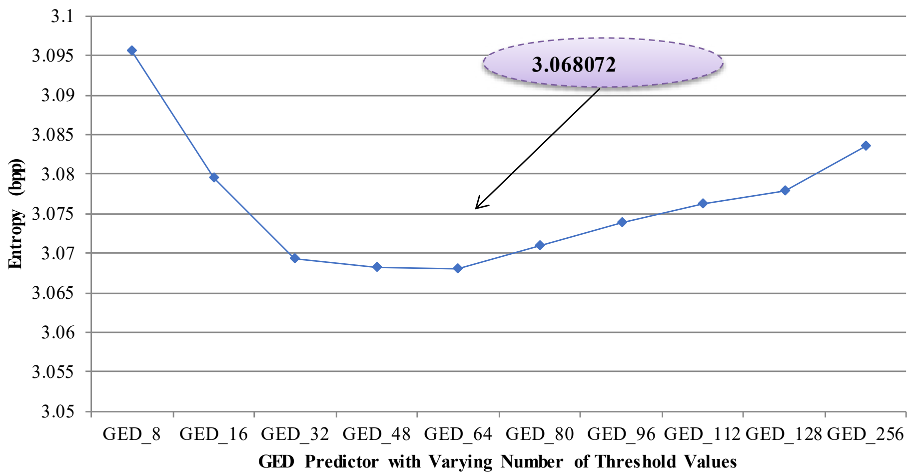


Figure 6. Average value of entropy from GED predictor with varying threshold values for R2 resolution.

Entropy for R3 Dataset for Varying Threshold Values. Some medical images may have different resolution other than 256×256 and 512×512. In experiment 3, proposed algorithm RIGED was applied on such images too. Weighted average of entropy for set of resolution other than 256×256 and 512×512 resolution was calculated. There is significant change in entropy value when threshold varied from 8 to 32 and then slight change in entropy was observed and a minimum of entropy was obtained at threshold value of 48. The result is graphically represented in Figure 7.

Calculation of entropy for all the modality with all the possible resolution for varying threshold values. To come with resolution independent GED, it is mandatory to compare all the modality with all the possible resolutions. Weighted average of entropy for the complete set is calculated and shown in Table 4.

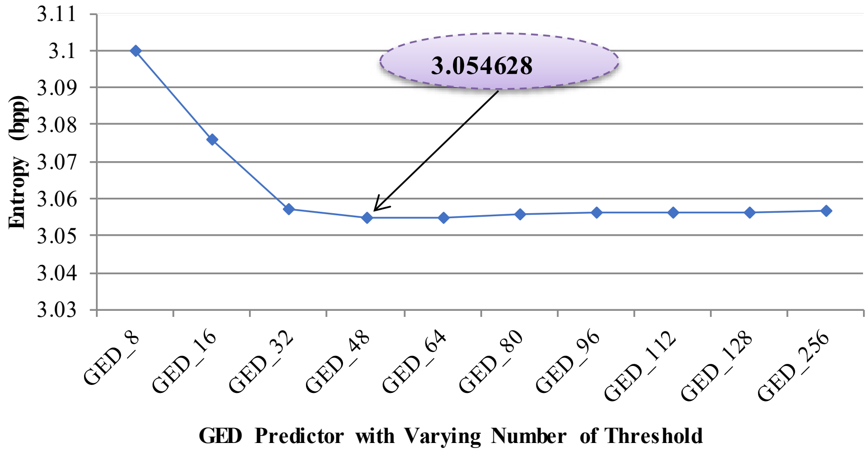


Figure 7. Average value of entropy from GED predictor with varying threshold values for image resolution in between 256×256 and 512×512

Table 4

Entropy values of medical dataset predicted from GED predictors with varying threshold values

Image Dataset Resolutions	No. of Slices	Entropy values before prediction	Entropy values after prediction									
			GED_8	**GED_16**	**GED_32**	**GED_48**	**GED_64**	**GED_80**	**GED_96**	**GED_112**	**GED_128**	**GED_256**
CT (256×256)	557	4.54622	2.69748	2.68278	2.67414	2.67442	2.67865	2.68250	2.68538	2.68732	2.68830	2.68977
CT (512×512)	322	4.67396	2.74261	2.72864	2.72852	2.72810	2.72270	2.728304	2.73388	2.73835	2.74138	2.75137
MRI (256×256)	305	5.31504	3.72402	3.71725	3.71375	3.71473	3.71702	3.71932	3.72096	3.72199	3.72253	3.72302
MRI (512×512)	211	5.27016	3.62855	3.59807	3.57815	3.57126	3.56793	3.56798	3.56806	3.56809	3.56807	3.56817
CT (336×336)	83	2.16590	1.45454	1.40205	1.36309	1.35401	1.35406	1.35409	1.35412	1.35421	1.35427	1.35438
SC (384×320)	29	6.20258	4.61906	4.60863	4.60216	4.60116	4.60413	4.60578	4.60660	4.60696	4.60709	4.60719
SC (320×270)	52	6.41378	5.32638	5.35222	5.36658	5.36136	5.37383	5.37521	5.37602	5.37651	5.37673	5.37709
XA (512×512)	90	5.83941	3.12632	3.12643	3.12950	3.12734	3.12574	3.13127	3.13209	3.13256	3.13300	3.13549
Weighted Average	1649	4.60039	3.08752	3.07197	3.06122	**3.05983**	3.06187	3.06468	3.06733	3.06938	3.07072	3.07483

Finally, a common threshold value was obtained for all the possible cases. Graphical representation of entropy from GED predictor at varying values of the threshold is shown in Figure 8. It was clear that GED with a threshold value of 48 gave a lower value of entropy for a complete set of medical images compared to other threshold values.

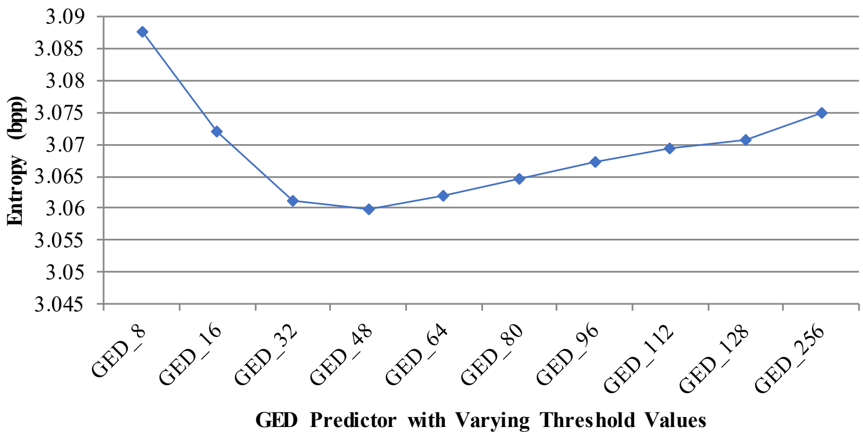


Figure 8. Average value of entropy from GED predictor with varying threshold values for overall dataset including different resolutions and different modalities.

For a complete set of medical images (Table 1, 2, 3), minimum entropy was in between the threshold 32 to 80. Weighted average of minimum entropy was obtained at threshold of 48 for all available dataset. To check the variation near to the optimal value and confirm our findings, entropy value was also calculated for threshold value before and after 48 for exact minima of entropy i.e., for threshold value of 43 to 53 in steps of 1 as represented in Figure 9. It is clear from the figure that entropy obtained at the value of threshold 44 (GED_44) was lower and then it again started increasing with increasing number of threshold value.

Result validation of prediction at optimal threshold values. From above discussion it is clear that lowest value of entropy was obtained for threshold range between threshold values 32 to 64 of GED predictor for complete set of medical images of R1, R2 and R3 resolution. GED with threshold 32 was optimal for 256×256 resolution and threshold of 64 gave optimized results of entropy for 512×512. Threshold of 44 was obtained as optimal for overall database of different modalities of medical images of varying resolution. There was a slight variation in entropy value between these thresholds. Results of medical images from RIGED predictor for all threshold values i.e., 32, 44, 48 and 64 are shown in Figure 10.

Comparison of RIGED with Standard MED and GAP Predictors in Terms of Entropy and Computational Complexity

Entropies of prediction error and computational complexity are two performance parameters used for evaluation of the proposed RIGED predictor. A predictor is efficient if it provides

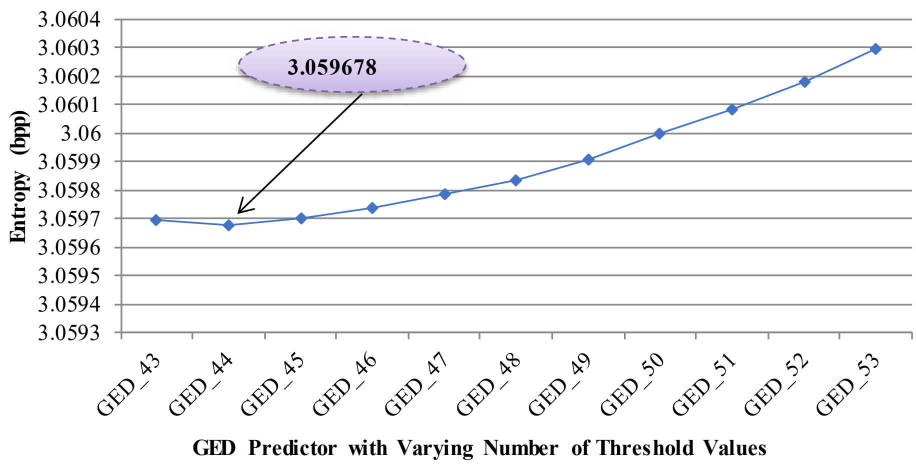


Figure 9. Average value of entropy from GED predictor with varying threshold before and after threshold point 48 for complete set of different modalities of medical images.

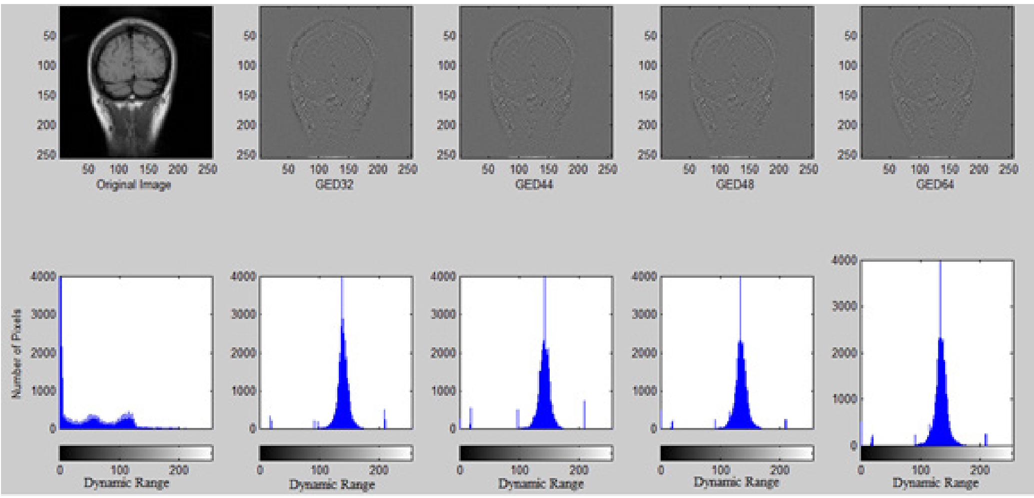


Figure 10. Original MRI-scan image of brain, residual image obtained from RIGED predictor with threshold values 32, 44, 48 and 64 and their histograms.

less entropy value with low computational complexity and run time. The lower the value of entropy obtained, the better is predictor's efficiency (Avramovic, 2012). Proposed algorithm gave better results than MED and comparable results as that of highly efficient standard GAP for same set of database. Comparison of RIGED done with state-of-the-art MED and GAP predictors in terms of entropy is shown in Table 5.

Table 5

Entropy values of medical dataset from different resolution and modality from different predictors and % improvement of RIGED over MED and GAP

Image Dataset Resolution	No. of Slices	Entropy values before prediction	Entropy values after prediction			% Improvement of RIGED	
			MED	**GAP**	**RIGED (Proposed)**	**Over MED**	**Over GAP**
CT (256×256)	557	4.54622	3.75772	2.73264	**2.67414**	40.520	2.187
CT (512×512)	322	4.67396	3.52607	2.76799	**2.72852**	29.229	1.446
MRI (256×256)	305	5.31504	4.69134	**3.54982**	3.71375	26.323	-4.414
MRI (512×512)	211	5.27016	4.84046	3.57121	**3.56793**	35.665	0.091
CT (336×336)	83	2.16590	1.92817	1.48989	**1.35611**	42.183	9.864
SC (384×320)	29	6.20258	5.64326	**4.34587**	4.60161	22.636	-5.557
SC (320×270)	52	6.41378	6.19173	**5.02089**	5.37052	15.291	-6.510
XA (512×512)	90	5.83941	4.18896	**2.8392**	3.12665	33.975	-9.193
Weighted Average	1649	4.60039	4.05145	3.03867	3.05967	32.414	-0.686

It is clear from Table 5 that minimum entropy is achieved by GAP predictor and RIGED has comparable results. Percentage improvement of RIGED over other predictors is also shown in this table and it is revealed that RIGED has a maximum improvement for CT image having resolution 336×336. Average percentage improvement of RIGED for complete sample dataset is 32% over MED and the percentage difference between GAP and RIGED is 0.68 % i.e., performance in terms of entropy for RIGED is 0.68 % less than GAP. RIGED is simple to implement as compared to GAP and also gave comparable results in terms of entropy. But at the overall new threshold value, RIGED performs better for some medical image datasets than highly efficient GAP. In literature entropy of prediction error image from different predictors is calculated only for a small set of medical images that are reported in Table 6. In this research work, weighted average of entropy is calculated for different modality with varying resolution and slices.

As GED is a threshold based and no specific threshold value is provided in literature. Threshold value of 16, 64, and 128 for GED prediction are randomly selected and reported by Avramovic and Savic (2011). These threshold values were tested on some MRI and CT image slices as shown in Table 6.

In this paper, we work on large set of volumetric medical database having varying number of slices, image resolution and modalities. For comparison purpose, reported

GED threshold values are tested on same set of dataset on which RIGED is implemented. Compared results between with RIGED and without RIGED is tabulated in Table 7.

Table 6

Entropy values of medical dataset from different predictors in literature

Images datasets		**Image Size**	**Slices**	**Entropies after prediction**					
				MED	GAP	GED*	GED 16	GED 64	**GED 128**
Avramovic and Savic (2011)	CT1	512×512	38	7.615	7.563	-	7.606	7.495	**7445**
	CT2	512×512	50	5.796	5.727	-	5.779	5.758	**5787**
	CT3	512×512	14	3.998	4.304	-	4.598	4.711	**4786**
	MRI1	640×576	25	4.747	4.851	-	5.072	5.264	**5359**
	MRI2	378×384	15	4.253	4.375	-	4.684	4.876	**4945**
Avramovic (2012)	CT	512×512	1338	4.19	4.39	4.30	-	-	-
	MRI	512×512	528	5.58	5.57	5.76	-	-	-

GED* threshold value is not specified.

GED16 indicates GED with threshold value 16.

GED64 indicates GED with threshold value 64.

GED128 indicates GED with threshold value 128.

Table 7

Comparison of with RIGED and without RIGED in terms of entropy value

Image Resolution	**Without RIGED**			**With RIGED**	**Percentage Improvement of with RIGED over without RIGED**		
	Threshold =16	Threshold =64	Threshold =128		Threshold =16	Threshold =64	**Threshold=128**
256×256	3.0488	3.0460	3.0542	**3.0419**	0.22	0.13	**0.40**
512×512	3.0795	3.0680	3.0779	**3.0680**	0.37	0.00	**0.32**
Other combinations of resolutions	3.0719	3.0618	3.0707	**3.0596**	0.40	0.07	**0.36**
Average Percentage Improvement					**0.33 %**	**0.07 %**	**0.36 %**

Average percentage improvement of with RIGED is around 0.33 % over without RIGED when threshold value is taken as 16. Proposed RIGED performs 0.07 % and 0.36 % better than without RIGED for threshold value of 64 and 128 respectively.

Complexity of RIGED was measured on the basis of arithmetic operations used for predictor's implementation and on the basis of run time of predictor. Total numbers of arithmetic operations involved in implementation of RIGED algorithm were calculated. A minimum of 8 and maximum of 11 operations were required for its implementation. A comparison of computation complexities in GAP, MED and RIGED is tabulated in Table

8. As observed, total number of minimum and maximum operations required for RIGED prediction was less as compared to highly efficient GAP predictor.

Table 8
Computational complexity of GAP and RIGED predictor

Predictors	Operations	Number of Addition/Subtraction	Number of Multiplication/Division	Number of Comparison	Total	Run Time* (μs)
MED	Minimum	-	-	1	1	2.51
	Maximum	2	-	2	4	
GAP	Minimum	11	-	1	12	5.38
	Maximum	18	8	6	32	
RIGED	Minimum	7	-	1	8	4.11
	Maximum	9	-	2	11	

*Tested on AMD A8-7410 APU with AMD Radeon R5 Graphics. 64 bit operating system and x64 based processor. Software: MATLAB 2013.

Time period was measured on the basis of time required for the execution of a predictor. MR-image sample of resolution 256×256 was tested on AMD A8-7410 APU with AMD Radeon R5 Graphics, 64 bit operating system and x64 based processor to measure the execution time (Software: MATLAB 2013). Runtime (execution time) in μs was calculated for all the techniques, and GAP had a high execution time as compare to RIGED predictor as tabulated in the Table 8.

CONCLUSION

Compression based on predictive coding performs well as it has simpler implementation and low computational complexity. 2D predictors GAP and GED predictor with different thresholds are implemented that are used for removing inter-pixel redundancy. It is revealed that prediction based on edge detection called GED is a good solution for prediction. Comparison of GED with GAP predictors is also made in this paper. The entropy of GED is a calculated for different threshold values and it is concluded that for a different resolution of image datasets, GED has optimum results at different thresholds.

The weighted average of entropy achieved by RIGED for different types of the database having varying resolution and it gave optimal threshold value at different image resolution. A common threshold value is selected that is optimal for every possible combination of resolutions and image modalities.

It is concluded that RIGED predictor can achieve comparable results in terms of entropy and it is also simple to implement as compared to GAP. It is revealed from the study that RIGED takes advantage of MED and GAP predictors as it is efficient and also simple to implement. Average percentage improvement of RIGED for complete sample dataset is 32.4% over MED and the percentage difference between GAP and RIGED is 0.68 %.

ACKNOWLEDGEMENT

Authors would like to thank the Jaypee University Waknaghat, Distt. Solan, India for supporting and providing help to this research.

REFERENCES

Al-Mahmood, H., & Al-Rubaye, Z. (2014). Lossless image compression based on predictive coding and bit plane slicing. *International Journal of Computer Applications*, *93*(1), 1-6.

Al-Naqeeb, A. B., & Nordin, M. J. (2017). Robustness watermarking authentication using hybridisation DWT-DCT and DWT-SVD. *Pertanika Journal of Science and Technology*, 25(S), 73-86.

Avramovic, A. (2012). On predictive-based lossless compression of images with higher bit depths. *Telfor Journal*, *4*(2), 122-127.

Avramovicl, A., & Savicl, S. (2011). Lossless predictive compression of medical images. *Serblan Journal of Electrical Engineering*, *8*(1), 27-36.

Avramovic, A., & Reljin, B. (2010, September 15-17). Gradient edge detection predictor for image lossless compression. In *Proceedings of the ELMAR* (pp. 131-134). Zadar, Croatia.

Baware, E. A., & Save, J. (2016). Medical image compression using adaptive prediction and block based entropy coding. *International Journal of Computer Applications*, *153*(9), 28-33.

Fouad, M. M. (2015). A lossless image compression using integer wavelet transform with a simplified median-edge detector algorithm. *International Journal of Engineering and Technology*, *15*(4), 68-73.

Hashemi-Berenjabad, S., & Mojarrad, S. (2016). A review on medical image compression techniques. International *Journal of Emerging Technologies in Engineering Research (IJETER)*, *4*(1), 88-91.

ISO/IEC 14495–1. (1999). *Information technology-lossless and near-lossless compression of continuous tone still images. JPEG-LS source code.* Retrieved August 29, 2019 from http://www.stat.columbia.edu/%7Ejakulin/jpeg-ls/mirror.htm.

Khatkar, K., & Kumar, D. (2018). A new genetic algorithm based technique for biomedical image enhancement. *Pertanika Journal of Science and Technology*, *26*(4), 1725-1750.

Kaur, H., Kaur, R., & Kumar, N., (2015). Review of various techniques for medical image compression. *International Journal of Computer Application*, *123*(4), 25-29.

Me, S. S., Vijayakuymar, V. R., & Anuja, R. (2012). A survey on various compression methods for medical images. *International Journal of Intelligent System and Applications*, *4*(3), 13-19.

Mofreh, A., & Refaat, A. M. (2016). A new lossless medical image compression technique using hybrid prediction model. *Signal Processing: An International Journal (SPIJ)*, *10*(3), 21-30.

Placidi, G. (2009). Adaptive compression algorithm from projections: Application on medical greyscale images. *Computers in Biology and Medicine*, *39*(11), 993-999.

Puthooran, E., Anand, R. S., & Mukherjee, S. (2013). Lossless compression of medical images using a dual level DPCM with context adaptive switching neural network predictor. *International Journal of Computational Intelligence Systems*, *6*(6), 1082-1093.

Sharma, U., Sood, M., & Bhardwaj, C. (2017). Reconstruction methods in compressive sensing for biomedical images. *Journal of Global Pharma Technology*, *6*, 134-143.

Shanmathi, G., & Maniyath, S. R. (2017). Comparative study of predictors used in lossless image compression. *Asian Journal of Applied Science and Technology*, *1*(5), 10-13.

Shrikhande, N. R., & Bairagi, V. K. (2013). Prediction based lossless medical image compression. International *Journal of Electronics and Communication Engineering and Technology*, *4*(2), 191-197.

Thombare, S. (2016). Low-complexity, lossless volumetric medical image compression with PPM-5 & linear. *International Journal of Innovative Research in Computer and Communication Engineering*, *4*(10), 17140-17144.

Weinberger, M. J., Seroussi, G., & Sapiro, G. (2000). The LOCO-I lossless image compression algorithm: principles and standardization into JPEGLS. I*EEE Transaction on Image Processing*, *9*(8), 1309-1324.

Wu, X., & Memon, N. (1996, May 9). CALIC a context based adaptive lossless image codec. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing* (Vol. 4, pp. 1890-1893). Atlanta, GA, USA.

Pertanika J. Sci. & Technol. 27 (4): 1693 - 1708 (2019)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Hybrid Feature Extraction and Machine Learning Approach for Fruits and Vegetable Classification

Nimratveer Kaur Bahia[1], Rajneesh Rani[1]*, Aman Kamboj[1] and Deepti Kakkar[2]

[1]*Department of Computer Science and Engineering, Dr. B. R. Ambedkar National Institute of Technology Jalandhar, Jalandhar, Punjab 144011, India*

[2]*Department of Electronic and Communication Engineering, Dr. B. R. Ambedkar National Institute of Technology Jalandhar, Jalandhar, Punjab 144011, India*

ABSTRACT

Manual fruits and vegetables detection become easy when it is done in small amount, but it is a tedious process and more labor is required when gigantic amount is considered. So, automatic detection of these comes into usage. This study took the images of fruits and vegetables as input to the very first stage of processing from where detection was done. The entire process constituted three stages: Background subtraction, extraction of color as well as texture features, and then classification. Background subtraction was performed using k mean clustering technique. Color features were identified using statistical features. To identify texture features Histogram of Oriented Gradient (HOG), Local Binary Pattern (LBP) and Gray Level Co-occurrence Matrix (GLOM) were used. For training and classification, Support Vector Machine (SVM) classifier had been used and performance of this classifier had been compared with K Nearest Neighbor (KNN) classifier. After comparing the results, it shows that accuracy of SVM was higher than that of KNN. The accuracy obtained by SVM with quadratic kernel function was 94.3%.

Keywords: Color, gray level co-occurrence matrix, K mean clustering, K nearest neighbor, support vector machine, texture

ARTICLE INFO

Article history:
Received: 23 October 2018
Accepted: 04 June 2019
Published: 21 October 2019

E-mail addresses:
nimratbahia@gmail.com (Nimratveer Kaur Bahia)
ranir@nitj.ac.in (Rajneesh Rani)
amankambojj@gmail.com (Aman Kamboj)
kakkard@nitj.ac.in (Deepti Kakkar)
* Corresponding author

INTRODUCTION

Detection system is a 'magnificent challenge' posed to the computer vision for attaining recognition of the near human levels. An object can be very well detected using image processing techniques. There exists abundance of techniques used for the objects

ISSN: 0128-7680
e-ISSN: 2231-8526

detection but still lots of improvements are required to achieve high accuracy. Content based image retrieval can be applied for the object detection. It detects the image's contents that include color, texture, shape and size (Rocha et al., 2008). There are many applications of this work when used in the form of a mobile application. It helps the cashier during the billing and useful for the kids to increase their learning power (Zawbaa et al., 2014). The detection of fruits as well as vegetables is useful at places like supermarkets where cost involved for fruits purchased by any customer can be done automatically. Fruits as well as vegetables recognition may be brought into use in computer visualization for the automatic categorization of fruits and vegetables from any random badge, containing various fruits. The aim of this section is to represent a comprehensive overview of research development activities in the field of automatic fruit and vegetable classification.

Seng and Mirisaee (2009) had used the combination of color, size and shape features with KNN classifier to classify fruits and vegetables. The experiments done by them were on only seven categories and accuracy obtained was 90%. Ninawe and Pandey (2014) studied the fruit detection completion system by doing the experiments on a dataset of six varieties of fruit which contained red apple, green banana, green guava, green melon, orange and watermelon. They used four features: color, shape, texture, size and then used their combination for obtaining the better results. They used the geometrical properties to calculate the area as well as perimeter and computed roundness and entropy values for extracting different features. The classification was done using the KNN classifier. This method has a drawback as shape of each fruit or vegetable is different so we cannot recognize it on the bases of shape and size as it decreases the accuracy. Arivazhagan et al. (2010) had tried to do recognition of fruits and vegetables using minimum distance classifier based upon the statistical and co-occurrence features derived from the Wavelet transformed sub-bands. They achieved an accuracy of 87% approximately by doing experiments on a database of about 2635 fruits from 15 different classes. In this method they used 50% of images as training to maintain the high accuracy by using less number of color and texture features. Zhang and Wu (2012) had proposed a novel classification method based on a multi-class kernel support vector machine (kSVM) with the desirable goal of accurate and fast classification of fruits. The experimental results demonstrated that the Max-Wins-Voting SVM with Gaussian Radial Basis kernel achieved the best classification accuracy of 88.2%. In this method they had used 60% of images as training to achieve the high accuracy, but results were more accurate if training and testing images were in equal proportion. Dubey and Jalal (2015) proposed the method that recognized various numbers of different fruits as well as vegetables. They had extracted different color along with texture features. Global color histogram (GCH), Color coherence vector (CCV), Color difference histogram (CDH) had been used to extract color features and Structure Element Histogram (SEH), Completed Local Binary pattern (CLTP), Local Binary Pattern

(LBP) and Local ternary pattern (LTP) had been used to extract the texture features. For classification, they had used Multiclass SVM classifier for training and testing purpose and achieved 93.84% accuracy by doing experiments on a dataset of 2312 images. In this they used different methods for feature extraction but some of the methods did not help gain the accuracy. Jhawar (2016) sorted the orange fruit by using a pattern recognition technique. They took the 160 images which were gathered from various locations in Vidarbha area of Maharashtra. Only four features were extracted using maturity level. For doing the classification, three techniques were used: one was Edited Multi seed Nearest Neighbor Technique, second was linear regression technique while the third was nearest prototype. All were based on the technique of pattern recognition. The maximum accuracy obtained was 97.98% by using linear Regression classifier. The techniques used in this paper gave high accuracy, but they used very few numbers of images so, there was a need to test the method on large dataset to get accurate results. Shukla and Desai (2016) proposed the model which used machine learning for the automated recognition of fruit. In this, they used the color, texture, shape features followed by their combination for obtaining the better results. Color Coherence Vector (CCV) was used to extract the color features and GLCM and LBP techniques were used for texture features extraction and some statistical features were used for shape analysis. Two classifiers were used: KNN and SVM for the classification and the results were compared. The best accuracy obtained was 91.3% with KNN whereas by using the SVM classifier the accuracy was 86.96%. The technique is tested on very few images. Moreover, the images contain only one category of fruit. So, it is easy to detect fruits and vegetables in the images. Moallem et al. (2017) conducted a study which showed the grading of the golden delicious apple. To do this, they used different techniques. Stem end detection, clays detection, primary defect segmentation, refinement of defect regions was used to do the segmentation. After the segmentation was complete, they detected the defected regions corresponding to an image. After this statistical, textural and geometric feature were extracted. On finding the feature vector, various classifiers are applied. SVM, KNN, and MLP (multi-layer perceptron) classifiers were used. After classifying they divided the fruit into healthy and detected part. After that, healthy parts were further classified into first rank and second rank. SVM Classifier outperformed the other two type of classifiers with recognition rate of 92.5% and 89.2% for two categories. In this study, the author used only one type of fruit, which did not match with other fruits and vegetables due to this the method had achieved good accuracy. Another study by Wang et al. (2018) was on comfortable footwear design for patients with diabetic conditions. The features like HSV and HOG and GLCM were extracted and fed to the Fuzzy Support Vector Machine (FSVM) for the training of diabetic plantar pressure images. Their proposed system obtained an accuracy of 84.3% which is much better than in comparisons to SVM and LSVM. Ansari and Ghrera (2018) suggested a novel intuitionistic fuzzy feature extraction method to extract the local texture, in this method the authors had incorporated a new intuitionistic fuzzy set theory for the representation of pattern in the images. Their proposed method also contributes to more than one bin in the distribution, which was used as a feature vector. The proposed approach had shown much better results over the local binary pattern.

From the study of literature, it has been concluded that there is requirement to explore more state-of-the-art feature extraction and classification techniques to detect fruits and vegetables as majority of the methods used in literature have been tested on small datasets and images were having one type of fruit. So, in this paper we have utilized texture, color and hybrid features (color + texture) of an image for fruits and vegetable classification and compared their results using SVM and KNN classifier. Our method is evaluated on a large dataset and each image in the dataset contains multiple instances of the fruit or vegetable. The rest of paper is planned as follows: Section 2 explains the proposed methodology and the dataset used is implicated in the proposed work. Section 3 describes the various results obtained including the dataset of raw images obtained for research and compares the results with earlier approaches. Section 4 gives the conclusion and some future direction.

MATIRAL AND METHODS

Dataset

The dataset of fruits and vegetables used in this presented work is the same as used by Dubey and Jalal (2015). It consisted of 15 categories: Spanish Pear (159), Asterix Potato (182), Cashew (210), Nectarine (247),Plum (264), Onion (75), Granny-smith Apple (155), Orange (103), Tahiti Lime (105), Kiwi (151), Fuji Apple (212), Watermelon (192), Diamond Peach (211), Agata Potato (201) and Honeydew Melon (145): total of 2612 images. All this data was collected from supermarket where there were different types and varieties of fruits and vegetables. Figure 1 represents the dataset of different kind of Fruits and Vegetables (Dataset is accessible at – http://www.ic.unicamp.br/~rocha/pub/downloads/tropical-fruits-DB-1024x768.tar.gz). The Dataset contains more than one fruit in one image.

Methodology

The methodology used in the present work is described in this section as shown in Figure 2. The set of raw images had been first passed through K mean clustering technique for background subtraction and then various color and texture features had been extracted from the image as a feature vector and SVM classifier had been trained for the classification purpose.

During the training phase, both raw images (fruits and vegetables) and their corresponding labels were fed to the classifier. At the time of testing the same procedure was followed and classifier returned the labels corresponding to the input image of fruit or vegetable.

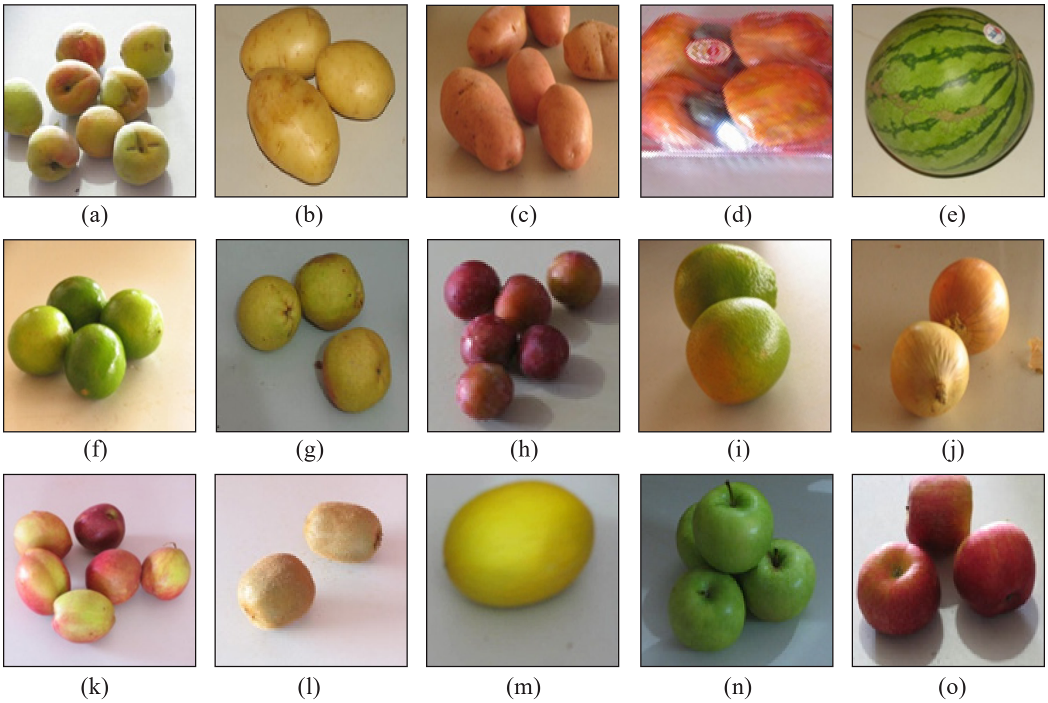


Figure 1. Data set of 15 different kinds of fruit and vegetable. (a) Diamond peach; (b) Agata Potato; (c) Asterix Potato; (d) Cashew; (e) Watermelon; (f) Tahiti Lime; (g) Spanish Pear; (h) Plum; (i) Oranges; (j) Onions; (k) Nectarine; (l) Kiwi; (m) Honeydew melon; (n) Granny Smith Apple; and (o) Fuji Apple

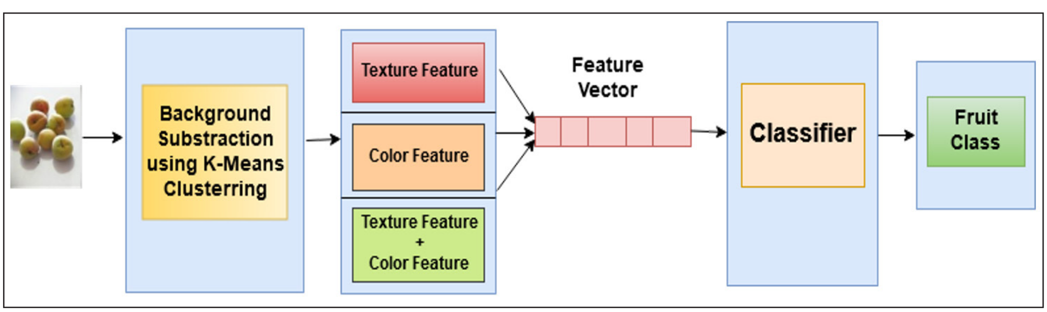


Figure 2. Flow chart of the proposed methodology

Background Subtraction. For performance enhancement of the proposed approach, supermarket dataset of 15 different categories of fruits along with vegetables was used. These images were used as the input. As the dataset used in the present work had been gathered from the supermarket so the data contained noisy or blurred images. For this type of raw data, background subtraction was made for extraction of the region of interest from the images. Background subtraction is the technique in which only foreground object is displayed and rest all another object which comes in the background is black. This will remove the background and extract foreground.

In the present work, for background subtraction, K-Mean Clustering technique had been used. The image was divided into k segments. K was taken as 2 because we had to divide the image into 2 segments in one part only the fruit or the vegetable would be shown and rest all the background part would be black. Figure 3 and Figure 4 show an example of K-Mean Clustering.

(a) (b) (c)

Figure 3. Before K-Mean Segmentation

(a) (b) (c)

Figure 4. After K-Mean Segmentation

The Figure 3 shows the images collected from supermarkets. In these figures, images have unwanted objects. In Figure 3(a) the image has a shadow of the fruit and in Figure 3(c) hand is captured in the image. Images after the preprocessing are displayed in Figure 4. After doing the k means clustering, we removed all the types of noise present in the images which got captured from the supermarket.

Feature Extraction. The data obtained after the background subtraction contains all the information that is required for extracting the desired results. The main motive of feature extraction is to acquire the most appropriate information from the data and constitute that information in a lower dimensionality space. When the input data to an algorithm is very large to be operated and it is imagined to be unnecessary then input data is turned into a few sets of features. Feature Extraction is the conversion of input images into different features suggested by (Zaitoun & Aqel, 2015). Many fruit detection and classification systems are based on color features along with shape features. But there are some types of fruits which have the samecolor and shape features. In the present work, extraction has been done by using two kinds of features: color and texture. Shape features are not used, because, in our dataset, there are more than one fruit, which is present in one image and cannot define its shape.

Color Feature Extraction. Color is the best feature to distinguish between fruits and vegetables. Because we see that almost all fruits and vegetables are of different colors, so we can easily distinguish. The various color spaces exist like RGB (Red, Green, and Blue) and HSI (Hue, Saturation and Intensity) and many more. Each color space has its own importance and provides color information in a more intuitive way. So, in the presented work both RGB and HSI color spaces had been used to extract the color features. After doing background subtraction image was in RGB color space. Six Color features had been extracted by using RGB color space: Kurtosis, standard deviation and skewness for the given RGB image and mean of all components that was mean of Red, Green and Blue component of the image. Three features were extracted after converting the image into HIS color space. Mean values of hue, saturation and intensity were calculated. Thus, a total of nine color features had been extracted.

Texture Feature Extraction. The texture is an important feature for detection of any object or to recognize some part from an image. A single pixel can never reveal the texture of the surface of an object. When there are a group of different pixels, then it becomes the texture element (Abdelmounaime and Dong-Chen, 2013; Lalibertea and Rango, 2008). The texture is calculated by the outer part of an object which measures the roughness, coarseness, and smoothness (Reddy et al., 2009). The neighborhood of an image is spatially distributed and specifies its texture (Pujari et al., 2013; Clausi, 2002). In this work, seven textural features had been extracted: five features such as contrast, correlation, homogeneity, energy, entropy with the help of GLCM while two other features were extracted with another two techniques HOG and LBP.

GLCM (Gray Level Co-occurrence Matrix). The textural feature is a type of gray-tone special dependencies which is helpful for the recognition of an image. The gray level matrix consists of a two-dimensional histogram, which is divided by a permanent spatial relationship. It is the statistical method that gives the spatial correlation of pixels. These can be calculated by first converting the RGB to gray-scale image (Sonka et al., 2014). In the proposed method first GLCM had been created, and then different statistics were calculated, which gave the information about the texture of an image. Different statistics used in this work are Entropy, Homogeneity, Contrast, Correlation, and Energy.

LBP (Local Binary Pattern). Prakasa (2016) worked with LBP technique for texture feature extraction. LBP extracts the surface of the image. In this, the texture pattern probability was computed and represented into histogram. They concluded the two type of LBP; one was symmetric and other was natural. For pattern classification, the LBP texture features could also be used. It is a very effective visual distributor technique for extracting the texture feature (Shukla & Desai, 2016). LBP is calculated by equating the adjacent pixels of an image (Dubey & Jalal, 2015). Ansari and Ghrera (2016) had proposed intuitionistic fuzzy local binary for extracting texture features from an image, suggesting extended fuzzy local binary pattern by incorporating intuitionistic fuzzy local binary pattern. The method was applied on various images and results indicated the effectiveness of the proposed method. The studies indicated that an LBP was a great texture feature used for categorization of the objects. In the present work, after extracting the LBP for each pixel, a histogram was created which represented the texture and mean of the histogram for an image was calculated and was labeled as LBP feature.

HOG (Histogram of oriented gradients). HOG utilizes overlapping local contrast normalization, while gets calculated on an evenly spaced cells' dense grid for enhanced accuracy (Dalal & Triggs, 2005). This texture feature had been calculated by taking the mean of luminance gradient component of each pixel of an image after converting it into gray scale image.

Classification

The process of classification with the help of feature vectors aids in the detection of fruits and vegetables. It defines boundaries between special targets in feature space with the help of extracted features as independent variables. Recent research has used a variety of machine learning models for example, KNN, SVM, decision trees and Neural Networks (NN) and their variants for this purpose. Linear and non-linear hyperdimensional data can be classified with the SVM which is a non-linear mapping of data with the help of kernel functions. KNN is an instance based non-parametric similarity measure learning for data of infinite dimensions and a decision tree is a probability-based graph for multi-class classification. SVM and KNN have been widely used for fruit and vegetable classification and a comparable classification effectiveness with respect to Multi-layer Perceptron (MLP) and Radial Bias Functions (RBF) has been reported (Hameed et al., 2018). That is why, in the present work, experiments had been done using Support Vector Machine (SVM) classifier and results had been compared with KNN classifier.

Support Vector Machine (SVM). The SVM is a machine learning tool for the use of data classification. SVM is used as the classification tool as it is a multiclass classifier with good accuracy and has got the ability to find a hyper plane with the widest margin that divides the samples into classes using kernel functions. There is one target value and various attributes for each instance in the training set (Prakash et al., 2012). The aim of the SVM is designing and producing such a model that will be able to predict the target value of data instances in the testing set where only attributes are being provided.

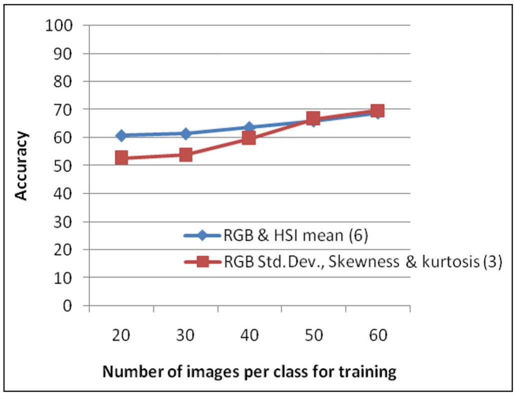
KNN (K Nearest Neighbor). KNN is the simplest classifier among all the other classifiers. It is the non-parametric method which is used for regression as well as for classifier. It does the classification on the basis of distance by measuring the distance matrix (Teoh et al., 2004). In this, K is the value used to make the boundaries of each and every class. When the value of K increases or decreases it affects the boundary values of class and error rate. When the value of K is 1 then the error rate is zero for the training sample. As the value of k increases, the boundary becomes smoother and error rate also increases.

RESULTS AND DISCUSSIONS

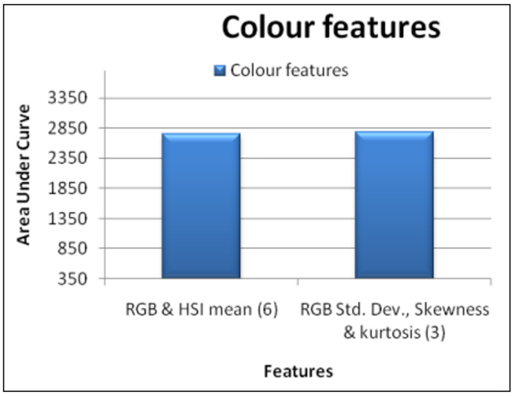
The study compares the various color features and texture features and then uses a combination of both, to find optimized features for the classification of fruits and vegetables. The experiments had been done using a different number of training images per class (20, 30, 40, 50 and 60) and the rest of the images had been taken as test images.

Analysis of Color Features

Figure 5(a) shows the average accuracy of fruits and vegetables using a different combination of color features. First, six color features which are the statistic mean a featurethat is red mean, blue mean, green mean, hue mean, saturation means, brightness mean. Second, a combination of color features: standard deviation, skewness, the kurtosis of RGB image has been used. Quadratic SVM is used to compute the accuracy.



(a)



(b)

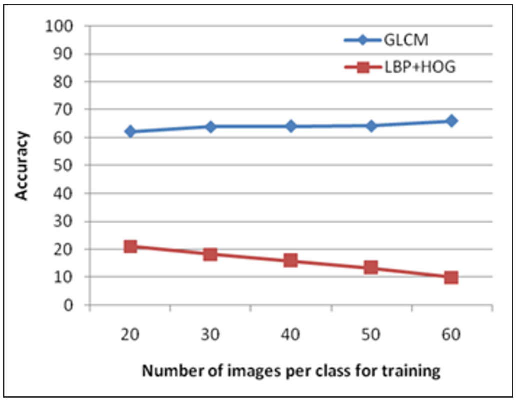
Figure 5. Average accuracy and area under curve of color features

Figure 5(b) shows the area under the curve for the same combination of color features. From Figure 5(a), accuracy increases as the number of training images per class are increased and maximum accuracy obtained is approximately 70%.

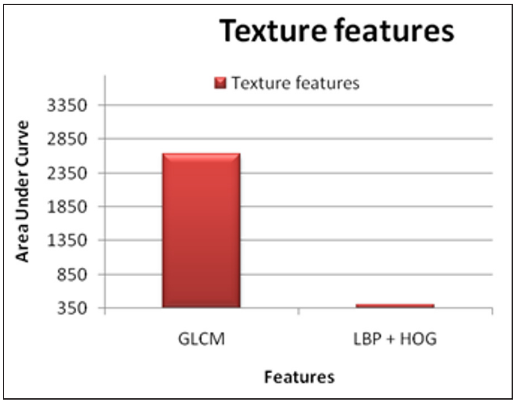
Analysis of Texture Features

The experiments had been done using a different combination of texture features. First, the combination of two texture features LBP and HOG and the second combination of five statistical GLCM features had been considered for experimentation. Statistical features of GLCM are energy, homogeneity, contrast, entropy, correlation. Average accuracy had been calculated by using Quadratic SVM and is represented in Figure 6(a). The area under the curve was also calculated for different features and is represented in Figure 6(b).

From the figures 6(a) and 6(b), it depicts that GLCM features perform very well than other two types of features in all cases.



(a)



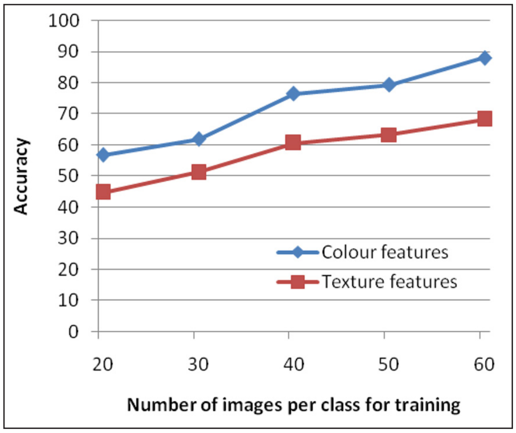
(b)

Figure 6. Average accuracy and area under curve of texture features

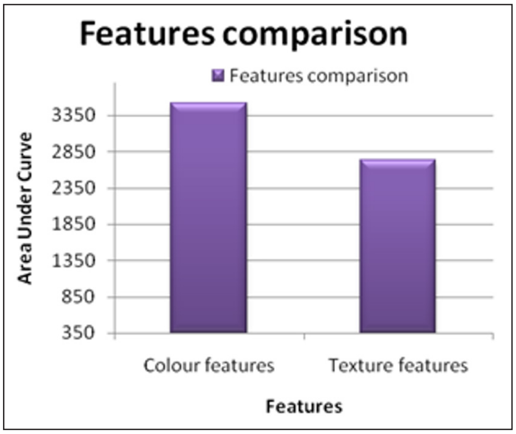
Analysis of Hybrid features (Combination of Texture and Color features)

Different experiments had been done by using different combinations of color and texture features. In the first experiment, all the color features together and all texture features together had been considered and computed the accuracy and area under the curve by using Quadratic SVM. Results are shown in Figure 7.

In the second experiment, a combination of some texture feature and some color feature had been used to find its corresponding accuracies. Combination of GLCM texture features with some color features which were statistical mean of RGB and HSI. HOG and LBP texture features were combined with some other color features which were skewness, standard deviation, and kurtosis of RGB image. Figure 8 shows the average accuracy by using a different number of images per class for training and area under the curve when we used color and texture features.

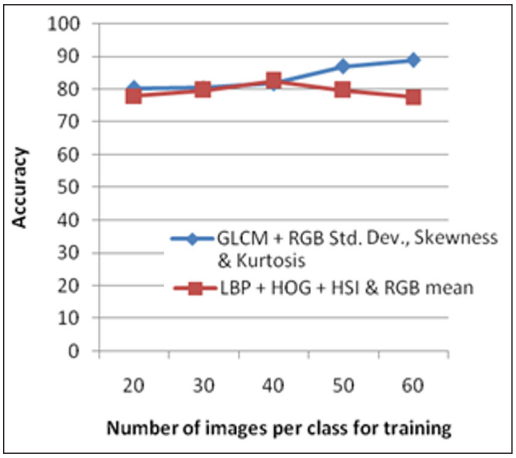


(a)

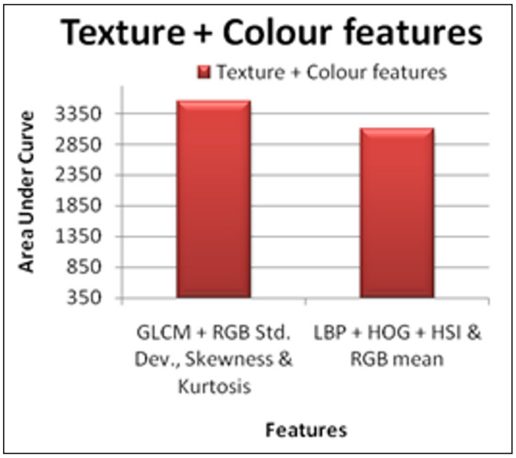


(b)

Figure 7. Average Accuracy and Area under curve of color and texture feature



(a)



(b)

Figure 8. Average Accuracy and area under curve when using combination of texture and color features

In the third experiment we evaluated all the sixteen features using SVM classifier to compare its results with KNN. The Table 1 shows the testing accuracy of SVM classifier with different kernel functions and Table 2 shows the testing accuracy of KNN classifier of its different types.

The comparison of KNN classifier with SVM classifier is shown in Figure 9. After comparing, the best result was generated with the SVM model with Quadratic kernel function and one-against-all strategy with a testing accuracy of 94.3%. While for KNN it gave the best accuracy of 74.66%. We conclude that SVM is better classifier to classify the fruits and vegetables than KNN classifier. This may be quadratic kernel function used in the SVM classifier which clearly separates the boundaries of different fruit and vegetables.

Table 1
Comparison of SVM and their corresponding accuracies

Sr. No.	Classifier	Accuracy
1.	Linear SVM	**80.19%**
2.	**Quadratic SVM**	**94.3%**
3.	Cubic SVM	**84.67%**
4.	Fine Gaussian SVM	**68.80%**
5.	Medium Gaussian SVM	**73.48%**
6.	Coarse Gaussian SVM	**52.45%**

Table 2
Comparison of KNN and their corresponding accuracies

Sr. No.	Classifier	Accuracy
1.	**Fine KNN**	**75.6%**
2.	Medium KNN	**68.98%**
3.	Coarse KNN	**48.89%**
4.	Cosine KNN	**70.38%**
5.	Cubic KNN	**68.28%**
6.	Weighted KNN	**74.59%**

In the fourth experiment, all nine color features with all the seven texture features have been considered and computed the average accuracy using Quadratic SVM Classifier. The test accuracy of each type of fruits and vegetables by using only color, only texture, and the combination of both, with Support Vector Machine, is mentioned in Table 3. The average accuracy has been found at 94.3%.

Table 3
Percentage of test accuracy of different fruits and vegetables

S. No.	Fruit or Vegetable name	Texture Features	Color Features	All Color + All Texture Features
1	Diamond Peach	71.75%	83.60%	**89.4%**
2	Agata Potato	56.42%	74.46%	**91.5%**
3	Cashew	81.33%	70.86%	**95.33%**
4	Fuji Apple	58.55%	89.33%	**90.8%**
5	Granny Smith Apple	72.63%	85.53%	**100%**
6	Honeydew Melon	97.64%	100%	**97.6%**
7	Kiwi	49.54%	92.94%	**91%**
8	Onion	66.66%	90.09%	**100%**
9	Orange	62.79%	100%	**97.7%**
10	Tahiti Lime	65.21%	86.04%	**85%**
11	Water melon	62.12%	71.73%	**98.4%**
12	Spanish Pear	63.63%	97.72%	**93%**

Table 3 *(continue)*

S. No.	Fruit or Vegetable name	Texture Features	Color Features	All Color + All Texture Features
13	Plum	83.33%	83.83%	**90.6%**
14	Asterix Potato	59.35%	89.21%	**92.6%**
15	Nectarine	68.85%	93.58%	**93.04%**
	Overall Accuracy	**68.22%**	**87.5%**	**94.3%**

Comparison with Existing Approaches

Results are also compared with other existing fruit or vegetable detection methods in Table 4 and concluded that the proposed method has achieved better accuracy. Dubey and Jalal (2015) performed fruit classification on same dataset (which was used in this research) and used different feature extraction techniques. Their method obtained 93% accuracy. Arivazhagan et al. (2010) also worked on the same dataset and they used different feature extraction techniques. For classification of fruits and vegetables minimum distance classifier was used and an accuracy of 87% was obtained.

The results are also compared with the approaches which work on a different dataset. In Zhang and Wu (2012), the color, shape and texture features are used with Kernel SVM and an accuracy of 88.2% was achieved. In Shukla and Desai (2016), similar type of features were used and with KNN classifier and an accuracy of 91.3% was achieved.

Table 4
Fruit recognition accuracy comparison with existing approaches

Reference	Fruit/ Vegetable (Dataset)	Features Extraction	Classifiers	Accuracy
Seng & Mirisaee (2009)	7 types of Fruits (50)	Mean of color image + shape features	K-Nearest Neighbors Algorithm	**90%**
Arivazhagan et al. (2010)	15types of fruits (2612)	Statistical features, cluster shade, cluster prominence	Minimum Distance classifier	**87%**
Zhang & Wu (2012)	18 types of fruits (1653)	Color features, texture and shape features	Kernel SVM	**88.2%**
Dubey & Jalal (2015)	15typesof fruits (2612)	GCH, CCV, CDH, SHE, LBP, LTP, CLBP	Multiclass SVM	**93.84%**
Shukla & Desai (2016)	9 types of fruits (115)	Color, texture, shape features	Multiclass SVM and KNN	**87% with SVM and 91.3% with KNN**
Proposed work	15 types of fruits (2612)	Statistical features, GLCM, LBP, HOG	Quadratic SVM (one-against-all strategy)	**94.3%**

Different classifiers were used to find accuracy. In our proposed method we have archieved an accuracy of 94.3, which is far better than existing methods. So, it is concluded that proposed work performs better than the existing approaches with less error rate.

CONCLUSION AND FUTURE WORK

Detection of fruits and vegetables is done manually but it becomes a difficult task when it is done in the supermarket or in industry. Time taken by a human to detect the fruits and vegetables becomes high so to reduce the time and to increase the accuracy, automatic detection of fruits and vegetables comes into existence. This research work is used to recognize fruits and vegetables into different categories based on different color and texture features. First a background subtraction method is used to extract the desire region of interest and various color and texture features are extracted. Then two classifiers SVM and KNN were brought into use for training and testing of the images using the features extracted at the previous stage. It is concluded that color gives the superior result than the texture but when the color of some fruit or vegetables are same then it becomes difficult to classify them, so texture features are used to differentiate the fruit or vegetable which have the same color. To further improve the classification accuracy, color and texture features are hybridized and results are computed using SVM and KNN classifier and it is concluded that SVM gives the highest accuracy up to 94.3%.

In this proposed work a single image has multiple fruits or vegetables but of the same type. So, in future the same work can be extended for classification of different fruit or vegetable present in a single image. As the fruits and vegetables are efficiently classified on color and texture basis but in the further work a mobile based application can be devolved which capture images of the fruit and can identify its class. To further improve the accuracy a deep learning model can also applied, and the same work can be also extended by adding more fruits and vegetables images in the database.

ACKNOWLEDGEMENT

The authors are thankful to acknowledge the support of Department of Computer Science & Engineering, National Institute of Technology Jalandhar for carrying out this research work. We thank our colleagues who provided insight and expertise that greatly assisted in the present work.

REFERENCES

Abdelmounaime, S., & Dong-Chen, H. (2013). New Brodatz-based image databases for grayscale color and multiband texture analysis. *ISRN Machine Vision, 2013*, 1-14.

Ansari, M. D., & Ghrera, S. P. (2016, November 25-27). Feature extraction method for digital images based on intuitionistic fuzzy local binary pattern. In *International Conference on System Modeling and Advancement in Research Trends (SMART)* (pp. 345-349). Moradabad, India.

Ansari, M. D., & Ghrera, S. P. (2018). Intuitionistic fuzzy local binary pattern for features extraction. *International Journal of Information and Communication Technology, 13*(1), 83-98.

Arivazhagan, S., Shebiah, R. N., Nidhyanandhan, S. S., & Ganesan, L. (2010). Fruit recognition using color and texture features. *Journal of Emerging Trends in Computing and Information Sciences*, *1*(2), 90-94.

Clausi, D. A. (2002). An analysis of co-occurrence texture statistics as a function of grey level quantization. *Canadian Journal of Remote Sensing, 28*(1), 45-62.

Dalal, N., & Triggs, B. (2005, June 20-25). Histograms of oriented gradients for human detection. In *Computer Society Conference on Computer Vision and Pattern RecognitionCVPR 2005* (pp. 886-893). San Diego, United States.

Dubey, S. R., & Jalal, A. S. (2015). Fruit and vegetable recognition by fusing colour and texture features of the image using machine learning. *International Journal of Applied Pattern Recognition, 2*(2), 160-181.

Hameed, K., Chai, D., & Rassau, A. (2018). A comprehensive review of fruit and vegetable classification techniques. *Image and Vision Computing*, *80*, 24-44.

Jhawar, J. (2016). Orange sorting by applying pattern recognition on colour image. *Procedia Computer Science*, *78*, 691-697.

Lalibertea, A. S., & Rango, A. (2008, August 5-8). Correlation of object-based texture measures at multiple scales in sub-decimeter resolution aerial photography. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* (pp.1-6). Calgary, Alberta, Canada.

Moallem, P., Serajoddin, A., & Pourghassem, H. (2017). Computer vision-based apple grading for golden delicious apples based on surface features*. Information Processing in Agriculture, 4*(1), 33-40.

Ninawe, P., & Pandey, S. (2014). A completion on fruit recognition system using k-nearest neighbors algorithm. *International Journal of Advanced Research in Computer Engineering and Technology*, *3*(7), 2352-2356.

Prakasa, E. (2016). Texture feature extraction by using local binary pattern*. INKOM Journal, 9*(2), 45-48.

Prakash, J. S., Vignesh, K. A., Ashok, C., & Adithyan, R. (2012, December 14-15). Multi class Support Vector Machines classifier for machine vision application. In *International Conference onMachine Vision and Image Processing (MVIP)* (pp. 197-199). Taipei, Taiwan.

Pujari, J. D., Yakkundimath, R., & Byadgi, A. S. (2013). Grading and classification of anthracnose fungal disease of fruits based on statistical texture features. *International Journal of Advanced Science and Technology, 52*(1), 121-132.

Reddy, B. R., Suresh, A., Mani, M. R., & Kumar, V. V. (2009). Classification of textures based on features extracted from preprocessing images on random windows*. International journal of advanced Science and technology, 9*, 9-18.

Rocha, A., Hauagge, D. C., Wainer, J., & Goldenstein, S. (2008, October 12-15). Automatic produce classification from images using color, texture and appearance cues. In *XXI Brazilian Symposium on Computer Graphics and Image Processing (SIBGRAPI'08)* (pp. 3-10). Campo Grande, Brazil.

Seng, W. C., & Mirisaee, S. H. (2009, August 5-7). A new method for fruits recognition system. In *International Conference on Electrical Engineering and Informatics (ICEEI'09)* (pp. 130-134). Selangor, Malaysia.

Shukla, D., & Desai, A. (2016, December 19-21). Recognition of fruits using hybrid features and machine learning. In *International Conference on Computing, Analytics and Security Trends (CAST)* (pp. 572-577). Pune, India.

Sonka, M., Hlavac, V., & Boyle, R. (2014). Image processing, analysis, and machine vision. *Cengage Learning, 9*(4), 55-58.

Teoh, A., Samad, S. A., & Hussain, A. (2004). Nearest neighborhood classifiers in a bimodal biometric verification system fusion decision scheme. *Journal of Research and Practice in Information Technology, 36*(1), 47-62.

Wang, C., Li, Z., Dey, N., Li, Z., Ashour, A. S., Fong, S. J., & Shi, F. (2018). Histogram of oriented gradient based plantar pressure image feature extraction and classification employing fuzzy support vector machine. *Journal of Medical Imaging and Health Informatics, 8*(4), 842-854.

Zaitoun, N. M., & Aqel, M. J. (2015). Survey on image segmentation techniques. *Procedia Computer Science, 65*, 797-806.

Zawbaa, H. M., Abbass, M., Hazman, M., & Hassenian, A. E. (2014, November 28-30). Automatic fruit image recognition system based on shape and color features. In *International Conference on Advanced Machine Learning Technologies and Applications* (pp. 278-290). Cairo, Egypt.

Zhang, Y., & Wu, L. (2012). Classification of Fruits Using Computer Vision and a Multiclass Support Vector Machine. *Sensors*, *12*(9), 12489-12505.

Pertanika J. Sci. & Technol. 27 (4): 1709 - 1722 (2019)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

A New Method to Construct 4-Cycle-Free Parity-Check Matrices for Regular LDPC Codes

Djamel Slimani*, Adda Ali-Pacha and Naima Hadj-Said

Laboratory of Coding and Security of Information, University of Sciences and Technology of Oran Mohamed Boudiaf, Algeria USTO-MB, PoBox 1505 El M'Naouer Oran 31000 Algeria

ABSTRACT

Low-Density Parity-Check (LDPC) codes are considered among the best error-correcting codes in use today. These codes can be defined by a sparse parity-check matrix H, which has a graphical representation as a Tanner graph. Several studies have shown that the existence of 4-cycles in the Tanner graph affects the performance of LDPC codes. In this paper, we propose a method which allows the construction of 4-cycle-free parity-check matrices. The main principles behind the proposed method are as follows: First, we choose a vector V which consists of w_c ones and L-w_c zeros, in such a way that the chosen vector allows us to construct a circulant matrix H_1 without 4-cycles. Second, we pass this matrix to the proposed algorithm to obtain a set of L-vectors. When any vector taken from this set is appended as a news column in the matrix H_1, we obtain an L×(L+1) matrix without 4-cycles. Next, we select those vectors that lead to a circulant matrix H_2 without 4-cycles. Finally, we can obtain an L×2L matrix H without 4-cycles by concatenating matrices H_1 and H_2.Simulation results confirm that the structure of the matrices constructed by the proposed method significantly reduces the encoding complexity. Though the performance of these matrices at higher signal-to-noise-ratios (SNRs) is not as good as those constructed by MacKay's method, they can be applied to practical

ARTICLE INFO

Article history:
Received: 2 November 2018
Accepted: 13 May 2019
Published: 21 October 2019

E-mail addresses:
djamel.slimani@univ-usto.dz, djmslimani@gmail.com (Djamel Slimani)
a.alipacha@gmail.com (Adda Ali Pacha)
naima.hadjsaid@univ-usto.dz (Naima Hadj-Said)
**Corresponding author*

ISSN: 0128-7680
e-ISSN: 2231-8526

communications because of being encoded in linear time with shift registers.

Keywords: 4-cycle-free matrices, circulant matrices, LDPC codes, Tanner graph

INTRODUCTION

In the early sixties Gallager (1962) proposed a new class of block codes called low-density parity-check (LDPC) codes. Since then, these codes have been included in a variety of different standards including WiMAX (IEEE 802.16e), WiFi (IEEE 802.11n) and 10 Gb/s Ethernet (802.3an). An LDPC code is characterized by its sparse parity-check matrix H, which consists of zeros and ones with less ones than zeros. After the rediscovery of LDPC codes by MacKay and Neal (1995) in the mid-nineties, they quickly became very famous as they demonstrated an excellent performance (MacKay, 1999). The parity-check matrix H has a graphical representation known as the Tanner graph (Tanner, 1981). This graph is a bipartite one consisting of two kinds of nodes: bit (variable) nodes that correspond to the columns of H and check nodes that correspond to its rows. A bit node is connected to a check node by an edge if the value of the intersection of the column and the row corresponding to these nodes is equal to 1. A cycle in a bipartite graph is a set of edges that forms a continuous path starting with a node and returning to the same node without going through an edge more than once. The number of these edges is called the length of the cycle, and the smallest length is the girth of the code. The Tanner graph contains neither cycles of length 2 nor cycles of odd length. Therefore, the girth of this graph is at least 4. The number of cycles in the Tanner graph is one of the parameters that affect the performance of LDPC codes, particularly short cycles (Wiberg, 1996), and more precisely the cycles of length 4 (Li et al., 2017; MacKay, 1999). In the belief propagation algorithm (BPA), the presence of cycles in a Tanner graph causes a loss of independence in the messages sent by the nodes of the graph (Bandi et al., 2011).Various methods for constructing LDPC codes without short cycles have been proposed, such as constructions based on array dispersion and masking (Xu et al., 2016), based on difference sets (Esmaeili & Javedankherad, 2012) and based on sub-matrix shifting (Fan & Xiao, 2006a). The structure of the matrices constructed by these methods increases the encoding complexity.

To solve this problem, we propose a method to construct 4-cycle-free parity-check matrices, for regular LDPC codes, with low encoding complexity. These matrices having w_r ones in each row and w_c ones in each column where w_r=2w_c. In this work, only the values of w_c, which are equal to 2, 3, 4 and 5, are taken into account.

MATERIALS AND METHODS

Regular LDPC Codes

An LDPC code is characterized by its parity-check matrix H. This code is called regular if its parity-check matrix contains a fixed number w_r of ones in each row and a fixed number w_c of ones in each column (Johnson, 2010, p. 38). The matrix H represents the parity-check matrix of a regular LDPC code.

$$H = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}$$

The Tanner graph for the parity-check matrix H is shown in the Figure 1, where the dotted lines represent a 6-cycle.Since there are no smaller cycles in this graph; its girth is 6.

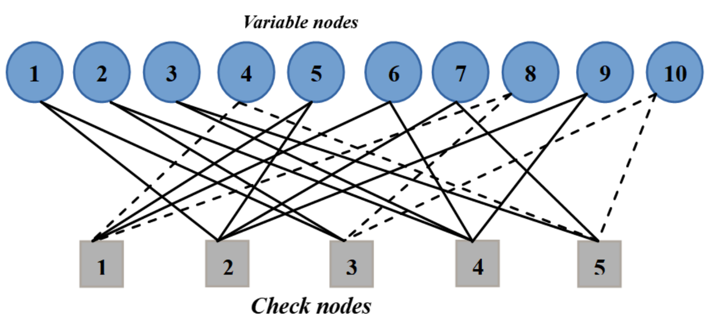


Figure 1. Tanner graph of the matrix H.

Circulant Matrices

A circulant matrix is a square matrix where each row is obtained by shifting the previous row one position to the right (Aldrovandi, 2001, p. 83), i.e,

$$H = \begin{pmatrix} h_0 & h_1 & h_2 & \cdots & h_{L-1} \\ h_{L-1} & h_0 & h_1 & \cdots & h_{L-2} \\ h_{L-2} & h_{L-1} & h_0 & \cdots & h_{L-3} \\ . & . & . & & . \\ . & . & . & & . \\ . & . & . & & . \\ h_1 & h_2 & h_3 & \cdots & h_0 \end{pmatrix}$$

In this paper, the values of h_i belong to the set $\{0,1\}$.

Construction of 4-Cycle-Free Matrices

The steps of the proposed method are as follows:

First step. In this step, we calculate all permutations P_{all} of the vector $V=[h_0\ h_1 h_2 \ldots h_{L-1}]$, which consists of *wc* ones and *L-wc* zeros, using the following law (Bóna, 2007, p.19):

$$P_{all} = \frac{L!}{w_c!(L-w_c)!} \qquad [1]$$

Using one of the cycle counting methods proposed by Fan and Xiao (2006b), Karimi and Banihashemi (2013), Li et al. (2015) or the permutations of the vector V which allow the construction of 4-cycle-free circulant matrices are found. We collect these permutations in a set S.

Based on simulation results, the length L is given by:

$$L \geq 2w_c\,(w_c - 1) + 1 \qquad [2]$$

For w_c equal to 2, 3 and 5, while for w_c equal to 4, L is given by:

$$L > 2w_c\,(w_c - 1) + 1 \qquad [3]$$

Second step. In this step, we took at each time a vector from the set S and constructed a circulant matrix that passed to the proposed algorithm detailed later on in the Proposed Algorithm section. The result of the algorithm is a set of all possible vectors $V_i' = [h_0' h_1' h_2' \ldots h_{L-1}']$ that allow us to use the matrix passed to the algorithm to construct an L×(L+1) matrix without 4-cycles as shown in the following matrix:

$$H = \begin{array}{c} L+1 \\ \left(\begin{array}{ccccc|c} h_0 & h_1 & h_2 & \cdots & h_{L-1} & h'_0 \\ h_{L-1} & h_0 & h_1 & \cdots & h_{L-2} & h'_1 \\ h_{L-2} & h_{L-1} & h_0 & \cdots & h_{L-3} & h'_2 \\ \cdot & \cdot & \cdot & \ddots & \cdot & \cdot \\ \cdot & \cdot & \cdot & \ddots & \cdot & \cdot \\ \cdot & \cdot & \cdot & \ddots & \cdot & \cdot \\ h_1 & h_2 & h_3 & \cdots & h_0 & h'_{L-1} \end{array}\right) \end{array} \; L$$

In general, not all vectors resulting from the algorithm allow the construction of circulant matrices without 4-cycles. For this reason, we chose only those that possessed the required characteristic of being able to construct matrices without 4-cycles. These vectors form a new set *S'*.

Third step. From the previous step, we noticed that each circulant matrix constructed by a vector of the set S had its own set *S'* which was found by using the proposed algorithm. In this last step, by concatenating each of these matrices with each circulant matrix, constructed by a vector of the set *S'*, we got matrices of the form:

$$H = \left(\begin{array}{ccccc|ccccc} \multicolumn{5}{c|}{L} & \multicolumn{5}{c}{L} \\ h_0 & h_1 & h_2 & \cdots & h_{L-1} & h'_0 & h'_1 & h'_2 & \cdots & h'_{L-1} \\ h_{L-1} & h_0 & h_1 & \cdots & h_{L-2} & h'_{L-1} & h'_0 & h'_1 & \cdots & h'_{L-2} \\ h_{L-2} & h_{L-1} & h_0 & \cdots & h_{L-3} & h'_{L-2} & h'_{L-1} & h'_0 & \cdots & h'_{L-3} \\ \cdot & \cdot & \cdot & \ddots & \cdot & \cdot & \cdot & \cdot & \ddots & \cdot \\ \cdot & \cdot & \cdot & \ddots & \cdot & \cdot & \cdot & \cdot & \ddots & \cdot \\ \cdot & \cdot & \cdot & \ddots & \cdot & \cdot & \cdot & \cdot & \ddots & \cdot \\ h_1 & h_2 & h_3 & \cdots & h_0 & h'_1 & h'_2 & h'_3 & \cdots & h'_0 \end{array}\right) \; L$$

The matrix H is an L×2L matrix without 4-cycles, which contains w_c ones in each column and $2w_c$ ones in each row.

$$S = \left\{\begin{matrix} 0000\mathbf{11}, 000\mathbf{1}0\mathbf{1}, 000\mathbf{11}0, 00\mathbf{1}0\mathbf{1}0, 00\mathbf{11}00, 0\mathbf{1}000\mathbf{1}, 0\mathbf{1}0\mathbf{1}00, 0\mathbf{11}000, \mathbf{1}0000\mathbf{1}, \\ \mathbf{1}000\mathbf{1}0, \mathbf{1}0\mathbf{1}000, \mathbf{11}0000 \end{matrix}\right\}.$$

Example 1. Let V be a vector of length L=6, which consists of 4 zeros and 2 ones.

Step 1. The number of permutations of V is P_{all}=15, only 12 permutations in the set S allow us to construct 4-cycle-free circulant matrices.

Step 2. We select a vector from the set S and construct a circulant matrix. This matrix is passed to the proposed algorithm. We find that each matrix passed to this algorithm has a set *S'* consisting of six vectors. For example; the circulant matrix constructed by the vector [100001] has the set:

$S'=\{\mathbf{1}0\mathbf{1}000,\mathbf{1}000\mathbf{1}0,0\mathbf{1}0\mathbf{1}00,0\mathbf{1}000\mathbf{1},00\mathbf{1}0\mathbf{1}0,000\mathbf{1}0\mathbf{1}\}$.

Step 3. For instance, we chose the vector [100001] with its own set *S'* (shown in the step 2) and we constructed the following 4-cycle-free circulant matrices:

$$H_1 = \left(\begin{array}{cccccc|cccccc} 1&0&0&0&0&1&1&0&1&0&0&0 \\ 1&1&0&0&0&0&0&1&0&1&0&0 \\ 0&1&1&0&0&0&0&0&1&0&1&0 \\ 0&0&1&1&0&0&0&0&0&1&0&1 \\ 0&0&0&1&1&0&1&0&0&0&1&0 \\ 0&0&0&0&1&1&0&1&0&0&0&1 \end{array}\right) H_2 = \left(\begin{array}{cccccc|cccccc} 1&0&0&0&0&1&1&0&0&0&1&0 \\ 1&1&0&0&0&0&0&1&0&0&0&1 \\ 0&1&1&0&0&0&1&0&1&0&0&0 \\ 0&0&1&1&0&0&0&1&0&1&0&0 \\ 0&0&0&1&1&0&0&0&1&0&1&0 \\ 0&0&0&0&1&1&0&0&0&1&0&1 \end{array}\right) H_3 = \left(\begin{array}{cccccc|cccccc} 1&0&0&0&0&1&0&1&0&1&0&0 \\ 1&1&0&0&0&0&0&0&1&0&1&0 \\ 0&1&1&0&0&0&0&0&0&1&0&1 \\ 0&0&1&1&0&0&1&0&0&0&1&0 \\ 0&0&0&1&1&0&0&1&0&0&0&1 \\ 0&0&0&0&1&1&1&0&1&0&0&0 \end{array}\right)$$

$$H_4 = \left(\begin{array}{cccccc|cccccc} 1&0&0&0&0&1&0&1&0&0&0&1 \\ 1&1&0&0&0&0&1&0&1&0&0&0 \\ 0&1&1&0&0&0&0&1&0&1&0&0 \\ 0&0&1&1&0&0&0&0&1&0&1&0 \\ 0&0&0&1&1&0&0&0&0&1&0&1 \\ 0&0&0&0&1&1&1&0&0&0&1&0 \end{array}\right) H_5 = \left(\begin{array}{cccccc|cccccc} 1&0&0&0&0&1&0&0&1&0&1&0 \\ 1&1&0&0&0&0&0&0&0&1&0&1 \\ 0&1&1&0&0&0&1&0&0&0&1&0 \\ 0&0&1&1&0&0&0&1&0&0&0&1 \\ 0&0&0&1&1&0&1&0&1&0&0&0 \\ 0&0&0&0&1&1&0&1&0&1&0&0 \end{array}\right) H_6 = \left(\begin{array}{cccccc|cccccc} 1&0&0&0&0&1&0&0&0&1&0&1 \\ 1&1&0&0&0&0&1&0&0&0&1&0 \\ 0&1&1&0&0&0&0&1&0&0&0&1 \\ 0&0&1&1&0&0&1&0&1&0&0&0 \\ 0&0&0&1&1&0&0&1&0&1&0&0 \\ 0&0&0&0&1&1&0&0&1&0&1&0 \end{array}\right)$$

To construct such matrices, it was enough to choose at each time a vector from the set S.Then, we passed the circulant matrix associated with this vector to the proposed algorithm so as to get a set of vectors *S'*. Matrices without 4-cycles would be constructed by concatenating the matrix passed to the proposed algorithm with each of the circulant matrices associated with the vectors of the set *S'*.

What about Large Matrices?

To construct large matrices, we simply added a number of zeros to the vectors obtained in first and second step in the Construction of 4-Cycle-Free Matrices section, in such a way that this number must be greater than or equal to L-1. These zeros had to be added, either on one side or on both sides of each vector as shown in the following example:

Example 2. The matrix H_2 of the example 1 consists of two circulant sub-matrices, the first sub-matrix is constructed by the vector [**1**0000**1**] while the second is constructed by the vector [**1**000**1**0]. To increase the size of H_2 without losing its characteristic (without

4-cycles), it is sufficient to add a number of zeros greater than or equal to 5 (because L=6) to each vector, either on one side or on both sides. To construct a 20×40 matrix, for example, we add to each vector 14 zeros. Then, the first vector becomes: [000000**100001**00000000] and the second becomes: [0000000000**10001**00000]. The matrix H_2 will be as follows:

$$H_2 = \left(\begin{array}{c|c}
00000010000100000000 & 00000000001000100000 \\
00000001000010000000 & 00000000000100010000 \\
00000000100001000000 & 00000000000010001000 \\
00000000010000100000 & 00000000000001000100 \\
00000000001000010000 & 00000000000000100010 \\
00000000000100001000 & 00000000000000010001 \\
00000000000010000100 & 10000000000000001000 \\
00000000000001000010 & 01000000000000000100 \\
00000000000000100001 & 00100000000000000010 \\
10000000000000010000 & 00010000000000000001 \\
01000000000000001000 & 10001000000000000000 \\
00100000000000000100 & 01000100000000000000 \\
00010000000000000010 & 00100010000000000000 \\
00001000000000000001 & 00010001000000000000 \\
10000100000000000000 & 00001000100000000000 \\
01000010000000000000 & 00000100010000000000 \\
00100001000000000000 & 00000010001000000000 \\
00010000100000000000 & 00000001000100000000 \\
00001000010000000000 & 00000000100010000000 \\
00000100001000000000 & 00000000010001000000
\end{array}\right)$$

The Proposed Algorithm

Here are some notations associated with the proposed algorithm:

I: a matrix which contains the indices of ones of the matrices *H* that will be passed to the proposed algorithm. Each row of *I* represents the indices of ones in each column of *H*. The following example shows the form of the matrix*I*:

$$H = \begin{pmatrix} 1\,0\,0\,1 \\ 1\,1\,0\,0 \\ 0\,1\,1\,0 \\ 0\,0\,1\,1 \end{pmatrix}, \qquad I = \begin{pmatrix} 1\,2 \\ 2\,3 \\ 3\,4 \\ 1\,4 \end{pmatrix}$$

For the first column of *H*, the indices of ones are '1' and '2' so we will fill the first row of *I* by 1 and 2 and so on until the last row of *I*.

L: the length of the vector used to construct the circulant matrix *H*.

V': defined in second step in the Construction of 4-Cycle-Free Matrices section.

Algorithm: Find a set S'

Inputs: H, L and w_c.
Output: Set S'.
1: Construct the matrix I, $v = [1\ 2\ 3\ ...\ L]$.
2: **for** $i = 1 \colon L$ **do**
3: Copy all rows of I that contain the value of i in a vector v_1 without repeating the value of i.
4: Extract the components that exist in the vector v and do not exist in the vector v_1, and replace the components of the vector v_1 by these extracted components.

Algorithm: Find a set S'

Update v_1 by deleting all components of this vector that are ***less than*** the value of i.
if $w_c = 2$ **then**
for $j = 1$: length of v_1 **do**
Initialize the vector V', of length L, to zero.
$V'(i) = V'\big(v_1(j)\big) = 1$.
endfor
else
for $j = 1$: length of v_1 do
Copy all rows of I that contain the value of $v_1(j)$ in a vector v_2 without repeating the value of $v_1(j)$.
Extract the components that exist in the vector v_1 and do not exist in the vector v_2, and replace the components of the vector v_2 by these extracted components.
Update v_2 by deleting all components of this vector that are ***greater than*** the value of $v_1(j)$.
if $w_c = 3$ **then**
for $k = 1$: length of v_2 **do**
Initialize the vector V', of length L, to zero.
$\mathrm{V}'(\mathrm{i}) = \mathrm{V}'\big(\mathrm{v}_1(\mathrm{j})\big) = \mathrm{V}'\big(\mathrm{v}_2(\mathrm{k})\big) = 1$.
endfor
else
for $k = 1$: length of v_2 **do**
Copy all rows of I that contain the value of $v_2(k)$ in a vector v_3 without repeating the value of $v_2(k)$.
Extract the components that exist in the vector v_2 and do not exist in the vector v_3, and replace the components of the vector v_3 by these extracted components.
Update v_3 by deleting all components of this vector that are ***greater than*** the value of $v_2(k)$.
if $w_c = 4$ **then**
for $l = 1$: length v_3 **do**
Initialize the vector V', of length L, to zero.
$V'(i) = V'\big(v_1(j)\big) = V'\big(v_2(k)\big) = V'\big(v_3(l)\big) = 1$.
endfor
elseif $w_c = 5$ **then**
for $l = 1$: length of v_3 **do**
Copy all rows of I that contain the value of $v_3(l)$ in a vector v_4, without repeating the value of $v_3(l)$.
Extract the components that exist in the vector v_3 and do not exist in the vector v_4, and replace the components of the vector v_4 by these extracted components.
Update v_4 by deleting all components of this vector that are ***greater than*** the value of $v_3(l)$.

Algorithm: Find a set S'

```
            for m = 1 : length v4 do
              Initialize the vector V', of length L, to zero.
              V'(i) = V'(v1(j)) = V'(v2(k)) = V'(v3(l)) = V'(v4(m)) = 1.
            endfor
          endfor
        endif
      endfor
    endif
  endfor
 endif
end for
```

RESULTS AND DISCUSSION

Results of the Proposed Method

In this section, we present the results of the proposed method. First, we calculated all permutations P_{all} of a vector V of length L.Then, we chose among these permutations only those that allowed the construction of matrices without 4-cycles (P_{free}). Finally, a number of 4-cycle-free matrices $H_{(L\times 2L)}$ was found according to the length L. These operations would be repeated for w_c equal to 2, 3, 4 and 5.

Table 1 shows the number of matrices $H_{(L\times 2L)}$ according to the length L, for w_c equal to 2. From the results in this table we note that the number of 4-cycle-free matrices increases proportionally to the length L. When L is odd, all permutations of V allow constructing matrices without 4-cycles, unlike if L is even. Tables 2, 3 and 4 show the number of matrices $H_{(L\times 2L)}$ for w_c equal to 3, 4 and 5, respectively, and for some values of L. The results in these tables show that the number of 4-cycle-free matrices varies according to the value of L, but not on a regular basis.

Table 1

Number of 4-cycle-free matrices for w_c=2

L	P_{all}	P_{free}	$H_{L\times 2L}$
≤4	/	/	0
5	10	10	50
6	15	12	72
7	21	21	294
8	28	24	384
9	36	36	972
10	45	40	1200
15	105	105	9450
20	190	180	28800

Table2
Number of 4-cycle-free matrices for w_c=3

L	P_{all}	P_{free}	$H_{L\times 2L}$
≤12	/	/	0
13	286	208	1352
14	364	196	0
15	455	360	12600
16	560	352	10240
17	680	544	46240
18	816	540	46656
19	969	798	138624
20	1140	800	137600

Table3
Number of 4-cycle-free matrices for w_c=4

L	P_{all}	P_{free}	$H_{L\times 2L}$
≤25	/	/	0
26	14950	4888	5408
27	17550	8748	104976
28	20475	7896	75264
29	23751	12180	518056
30	27405	11040	590400

Table4
Number of 4-cycle-free matrices for w_c=5

L	P_{all}	P_{free}	$H_{L\times 2L}$
≤40	/	/	0
41	749398	165968	26896
42	850668	140280	0
43	962598	234780	310632

Encoding Complexity Comparison

In this subsection, we compared encoding complexity of the proposed method with the methods proposed by Esmaeili and Javedankherad (2012); Fan and Xiao (2006a); MacKay (2005) and Xu et al.(2016). For that we used the encoding method proposed by Dutta and Pramanik (2015) to calculate the number of permutations needed to permute the parity-check matrix into the approximate UPPER triangular format (AUT). These permutations would be saved in a vector in order to apply the inverse permutation to each code word before it was transmitted. It is important to note that the encoding complexity increases as the number of permutations increases.

The 4-cycle-free parity-check matrices used in this comparison are as follows:

1. $V_1(9,18,26,28,32) = V_2(1,6,19,30,31) = 1$ (Size:102×204);
2. $V_1(16,17,19,25) = V_2(1,6,13,17) = 1$ (Size:124×248);
3. $V_1(12,16,18) = V_2(1,6,14) = 1$ (Size:189×378);
4. $V_1(22,31,39,41,45) = V_2(1,2,13,26,31) = 1$ (Size:408×816);
5. $V_1(16,18,19) = V_2(1,7,16) = 1$ (Size:432×864);
6. $V_1(1,6,10) = V_2(1,9,15) = 1$ (Size:765×1530);
7. $V_1(15,21,23,24) = V_2(1,6,16,20) = 1$ (Size:1020×2040);
8. $V_1(13,17,20) = V_2(1,2,19) = 1$ (Size:1080×2160);
9. $V_1(17,19,20) = V_2(1,6,20) = 1$ (Size:1320×2640);
10. $V_1(15,19,20) = V_2(1,10,18) = 1$ (Size:2000×4000);
11. $V_1(17,23,25,26) = V_2(1,5,12,17) = 1$ (Size:4000×8000);
12. $V_1(5,8,14) = V_2(1,3,16) = 1$ (Size:4320×8640);
13. $V_1(1,6,10) = V_2(1,11,14) = 1$ (Size:8000×16000);
14. $V_1(1,7,12) = V_2(1,10,13) = 1$ (Size:8640×17280).

For the proposed method:

For the method proposed by MacKay (2005):

1. 204.55.187(Size:102×204);
2. 816.3.174(Size:408×816);
3. 4000.2000.3.243 (Size:2000×4000);
4. 8000.4000.4.484 (Size:4000×8000).

For the method proposed by Xu et al. (2016):

By using the exponent matrix and masking matrix below, we constructed the following matrices by changing the identity matrix sizes:

1. Size of the identity matrix:330(Size of the resultant matrix:1320×2640);
2. Size of the identity matrix:500(Size of the resultant matrix: 2000×4000);
3. Size of the identity matrix:1000(Size of the resultant matrix: 4000×8000);
4. Size of the identity matrix:2000(Size of the resultant matrix:8000×16000).

$$P = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 23 & 67 & 83 & 117 & 142 & 158 & 206 & 249 \\ 46 & 134 & 166 & 234 & 284 & 316 & 82 & 168 \\ 69 & 201 & 249 & 21 & 96 & 144 & 288 & 87 \end{bmatrix}; \; M = \begin{bmatrix} 0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 & 1 & 1 \end{bmatrix}$$

For the method proposed by Esmaeili and Javedankherad (2012):

1. Difference set:(31,6,1),F32 : (Size of the resultant matrix:124×248);
2. Difference set:(31,6,1),F64 : (Size of the resultant matrix:189×378);
3. Difference set:(133,12,1),F256 : (Size of the resultant matrix:765×1530);
4. Difference set:(133,12,1),F256 : (Size of the resultant matrix:1020×2040);

For the method proposed by Fan and Xiao (2006a):

1. v=6,p=4 : (Size of the resultant matrix:432×864);
2. v=6,p=10 : (Size of the resultant matrix:1080×2160);
3. v=6,p=40 : (Size of the resultant matrix:4320×8640);
4. v=6,p=80 : (Size of the resultant matrix:8640×17280);

Table5

Comparison between the proposed and the first method

Methods	Matrices sizes			
	102 × 204	408 × 816	2000 × 4000	4000 × 8000
Proposed	14	07	09	09
Mackay (2005)	30	144	1900	3509

Table6

Comparison between the proposed and the second method

Methods	Matrices sizes			
	1320 × 2640	2000 × 4000	4000 × 8000	8000 × 16000
Proposed	08	09	09	10
Xu, Feng, Luo and Bai (2016)	704	1175	2321	4690

Table7

Comparison between the proposed and the third method

Methods	Matrices sizes			
	124 × 248	189 × 378	765 × 1530	1020 × 2040
Proposed	06	08	00	10
Esmaeili and Javedankherad (2012)	110	169	714	988

Table8

Comparison between the proposed and the fourth method

Methods	Matrices sizes			
	432 × 864	1080 × 2160	4320 × 8640	8640 × 17280
Proposed	06	11	09	12
Fan and Xiao (2006a)	402	1024	4213	8482

Tables 5, 6, 7 and 8 show some 4-cycle free parity-check matrices, constructed by the proposed method and the four methods mentioned above, and their corresponding number of permutations required to bring these matrices into the approximate upper triangular format. From the results shown in these tables, we can observe that the proposed method offered a gain of 16 permutations in the 102×204 matrix as a minimum gain, and 8470 as a maximum in the 8640×17280 matrix. For the 765×1530 matrix constructed by the proposed method, the number of permutations was zero, which meant that the associated encoder did not need to apply the inverse permutation to each code word before the transmission.

The proposed method gives the smallest number of permutations among the other methods for all matrices; this would imply that it offers the lowest encoding complexity. Furthermore, it provides more flexibility to construct matrices of different sizes than the other methods.

Performance of the Proposed Method

In this subsection we compared the performance, in terms of the bit error rate (BER), of two LDPC codes with parity-check matrices constructed by the proposed method and two LDPC codes with parity check-matrices constructed by MacKay's method (MacKay, 2005).The matrices constructed by MacKay are as follows:

1. 96.33.966 (w_c=3,w_r=6, size:48×96).
2. 96.44.443 (w_c=4,w_r=8, size:48×96).

Each of the matrices, constructed by the proposed method, is a concatenation of two matrices H_1 and H_2, where H_1 and H_2 are the circulant matrices constructed by the vectors V_1 and V_2 respectively. The matrices used in the simulation are defined by:

1. V_1 (1,7,12)=V_2 (1,10,13)=1 (w_c=3,w_r=6, size:48×96).

2. V_1 (17,23,25,26)=V_2 (1,5,12,17)=1 (w_c=4,w_r=8, size:48×96).

The encoding process can be realized using one of the methods proposed by Dutta and Pramanik (2015) or Richardson and Urbanke (2001). The encoded bits were modulated using binary phase shift keying ***(BPSK)*** before being sent over the additive white Gaussian noise ***(AWGN)*** channel. We used the ***sum-product*** algorithm (Johnson, 2010) for the decoding with a maximum of 10 iterations.

Figure 2 provides a performance comparison between the proposed method and MacKay's method for w_c=3 and 4. According to this figure, the proposed method slightly

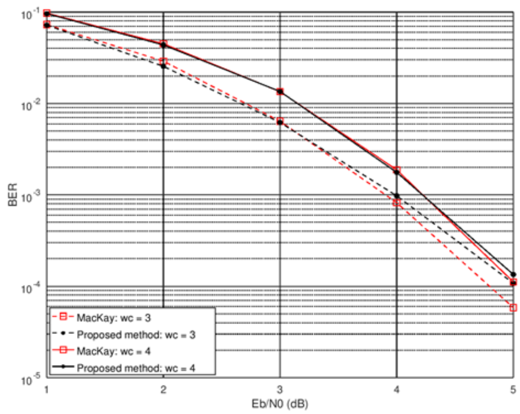


Figure 2. Performance comparison of the proposed method and that of MacKay

outperforms MacKay's method at lower SNRs, while at higher SNRs, MacKay's method performs better than the proposed method.

CONCLUSION

In this paper, we propose an original method capable of constructing 4-cycle-free parity-check matrices that can be used with regular LDPC codes. These matrices have full rank for odd-column-weight (3 and 5) and just one redundant row for even-column-weight (2 and 4). Simulation results show that the proposed method offers both an enormous reduction in encoding complexity and a very large number of 4-cycle-free parity-check matrices of different sizes. Although MacKay's method outperforms, in terms of BER, the proposed method in the high SNR region, the matrices constructed by the proposed method can be adopted in many practical applications due to their hardware implementation using simple shift registers.

ACKNOWLEDGEMENT

The first author would like to thank Dr. Abdellah Kaddai; Department of Electronics, Hassiba Ben bouali University at Chlef, Algeria; and Dr. Samir Bendoukha; Electrical Engineering Department College of Engineering at Yanbu Taibah University, Saudi Arabia; for their support and encouragement.

REFERENCES

Aldrovandi, R. (2001). *Special matrices of mathematical physics: Stochastic, circulant and bell matrices.* Singapore: World Scientific.

Bandi, S., Tralli, V., Conti, A., &Nonato, M. (2011). On girth conditioning for low-density parity-check codes. *IEEE Transactions on Communications, 59*(2), 357-362.

Bóna, M. (2007). *Introduction to enumerative combinatorics.* New York, NY: McGraw-Hill.

Dutta, A., & Pramanik, A. (2015, March 19-20). Modified approximate lower triangular encoding of LDPC codes. In *International Conference on Advances in Computer Engineering and Applications* (pp. 364-369). Ghaziabad, India.

Esmaeili, M., & Javedankherad, M. (2012). 4-Cycle free LDPC codes based on difference sets. *IEEE Transactions on Communications, 60*(12), 3579-3586.

Fan, J., & Xiao, Y. (2006a, November 6-9). A design of LDPC codes with large girth based on the sub-matrix shifting. In *IET International Conference on Wireless, Mobile and Multimedia Networks* (pp. 1-4). Hang Zhou, China.

Fan, J., & Xiao, Y. (2006b, November 16-20). A method of counting the number of cycles in LDPC codes. In *8th International Conference on Signal Processing* (pp. 2183-2186). Beijing, China.

Gallager, R. G. (1962). Low-density parity-check codes. *IRE Transactions on Information Theory, 8*(1), 21-28.

Johnson, S. J. (2010). *Iterative error correction: Turbo, low-density parity-check and repeat–accumulate codes.* New York, NY: Cambridge University Press.

Karimi, M., & Banihashemi, A. H. (2013). Message-passing algorithms for counting short cycles in a graph. I*EEE Transactions on Communications, 61*(2), 485-495.

Li, J., Lin, S., & Abdel-Ghaffar, K. (2015, June 14-19). Improved message-passing algorithm for counting short cycles in bipartite graphs. In *IEEE International Symposium on Information Theory* (pp. 416-420). Hong Kong, China.

Li, J., Lin, S., Abdel-Ghaffar, K., Ryan, W., & Costello, D. J. (2017). *LDPC code designs, constructions and unification.* New York, NY: Cambridge University Press.

MacKay, D. J. C. (1999). Good error-correcting codes based on very sparse matrices. *IEEE Transactions on Information Theory, 45*(2), 399-431.

MacKay, D. J. C. (2005). *Encyclopedia of sparse graph codes.* Retrieved April 11, 2002, from http://www.inference.phy.cam.ac.uk/mackay/codes/data.html.

MacKay, D. J. C., & Neal, R. M. (1995). Good codes based on very sparse matrices. In 5*th IMA Conference on Cryptography and Coding* (pp. 100-111). Berlin, Germany: Springer.

Richardson, T. J., & Urbanke, R. L. (2001). Efficient encoding of low-density parity-check codes. *IEEE Transactions on Information Theory, 47*(2), 638-656.

Tanner, R. M. (1981). A recursive approach to low complexity codes. I*EEE Transactions on Information Theory, 27*(5), 533-547.

Wiberg, N. (1996). *Codes and Decoding on General Graphs* (PhD dissertation). University of Linköping, Sweden.

Xu, H., Feng, D., Luo, R., &Bai, B. (2016). Construction of quasi-cyclic LDPC Codes via masking with successive cycle elimination. *IEEE Communications Letters, 20*(12), 2370-2373.

Pertanika J. Sci. & Technol. 27 (4): 1723 - 1733 (2019)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Matching Fingerprint Images for Biometric Authentication using Convolutional Neural Networks

Abdulmawla Najih*, Syed Abdul Rahman Al-Haddad Syed Mohamed, Abdul Rahman Ramli Shaiful Jahari Hashim and Nabila Albannai

Department of Computer and Communication Systems Engineering, Faculty of Engineering, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia

ABSTRACT

The use of biometric features, to authenticate users of different applications, is growing rapidly in recent years, according to the high sensitivity of the protected information and the good security that biometric authentication provides. In this study, a method is proposed to measure the similarity between two fingerprint images, using convolutional neural networks, instead of classifying them. Thus, modifying the users that the proposed method can recognize is a matter of adding or removing model images of the users' fingerprints. The similarity between the fingerprint image and every model image was measured in order to select the user with the highest similarity to the input image as the recognized user, where that similarity measure was compared to a threshold value in order to authenticate that user. The evaluation results of the proposed method, using FVC2002_DB1 and FVC2004_DB1 showed that the proposed method had 99.97% accuracy with 0.035% False Acceptance Rate (FAR) and 0% False Rejection Rate (FRR). Hence, the proposed method has been able to maintain high accuracy while eliminating the vulnerabilities of biometric authentication systems imposed by the use of separate stages for features extraction and similarity measurement.

Keywords: Biometric authentication, convolutional neural networks, fingerprints, machine learning

ARTICLE INFO

Article history:
Received: 29 November 2018
Accepted: 25 June 2019
Published: 21 October 2019

E-mail addresses:
nabdulmawla@gmail.com (Abdulmawla Najih)
sar@upm.edu.my (Syed Abdul Rahman Al-Haddad Syed Mohamed)
arr@upm.edu.my (Abdul Rahman Ramli)
sjh@upm.edu.my (Shaiful Jahari Hashim)
rehamranemrahf2009@gmail.com (Nabila Albannai)
* Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

INTRODUCTION

The rapid growth in Information Technology (IT) has emerged the need to protect sensitive and personal data from any unauthorized access. Many techniques are proposed to protect these data, such as the secret-based method, where login credentials are required from the users to access these data.

However, the importance of securing these data and the sensitivity of such methods to simple attacks, such as shoulder surfing, have imposed the need for more secure techniques (Nagatomo et al., 2018; Sun et al., 2018). Thus, the use of biometric authentication systems has attracted significant attention in recent years.

Biometric authentication systems rely on collecting distinctive information from a specific part of the human body, in order to distinguish one individual from another. The recognized individual can, then, be authenticated to the system or data protected by the biometric authentication system(McAteer et al., 2019), if that individual has the required privileges. This information can be extracted from the physiological or behavioral characteristics of the individual. These characteristics are evaluated using five quality measures, which are the acceptability, accessibility, availability, robustness and distinctiveness(Najih et al., 2016). Acceptable characteristics are those that can be collected from individuals without objections from them, according to some concerns such as privacy and security. The accessibility indicates the easiness of extracting this information from the individuals. Availability measures the ratio of individuals that these characteristics can be extracted from, with respect to the population. Robustness indicates the capability of extracting the same characteristics, every time this information is extracted from the individual, while the distinctiveness measures the variation in these characteristics among different individuals (Sinha & Ajmera, 2019; Zou et al., 2018).

Fingerprints are defined as the patterns created by the ripples in the skin of the human fingers. These patterns are very distinctive, where each human has different fingerprint, and very robust, as they do not change over time or because of any external conditions, such as wounds and scratches, where the same pattern is restored. Moreover, fingerprints are highly available in most humans, and do not threaten the privacy of individuals, hence, highly acceptable. Fingerprints can also be collected using cheap sensors that scan the fingers and extract their patterns (Alotaibi & Mahmmod, 2015; Douglas et al., 2018). However, some concerns have been shown regarding using a common surface to collect the fingerprint, which can participate in germs transportation from one individual to another. Thus, some touchless sensors have been implemented to eliminate such concerns, as well as, detecting vital signs from the finger, to deny the use of fake fingerprints (Orrù et al., 2019; Wang et al., 2016). These characteristics have encouraged the use of fingerprints in biometric authentication, rather than many other features, such as the face and iris. The use of facial features has risen some privacy concerns, while collecting information from the iris requires expensive equipment (Barni et al., 2015).

Many biometric authentication systems have been proposed based on fingerprints, where the individual is recognized based on the patterns collected from the fingerprint. Most of these systems measure the similarity between the collected fingerprint and those of the individuals that have the required privileges to access the system of information protected by the biometric authentication system. Different techniques are used to measure the similarity between these fingerprints. The biometric authentication system proposed by Kumar et al., (2016) used Speeded-Up Robust Features (SURF) descriptors to measure the similarity between fingerprint images. Despite the good performance of this method, which has shown only 0.06% EER using fingerprint images selected from the FVC2002 (Maio, Maltoni, Cappelli, Wayman, & Jain, 2002) and FVC2004 (Maio et al., 2004) datasets, the use of separate features extraction and matching stages imposes vulnerabilities to the biometric authentication system. Attackers may produce false features or tamper with the descriptors generated for the features before being matched at the matching stage (Ratha et al., 2001).

According to the outstanding performance of the artificial neural networks, on both accuracy and execution time measures, these networks have been employed to accelerate the performance of fingerprint biometric authentication systems. The method proposed by Peralta et al., (2018) uses a convolutional neural network to classify the fingerprint image into one of the five classes of fingerprints, defined by Henry, (1905) and shown in Figure 1. By classifying the input fingerprint images, as well as all the images of the known individuals in the database, the comparisons conduct to recognize the individual is limited to the number of model images that belong to the same class that the input fingerprint image belongs to. The performance of the convolutional neural networks is evaluated and compared to different other classifiers, using multiple datasets. The results show that the convolutional neural networks have outperformed all other classifiers in all of the used datasets, with a maximum classification accuracy of 99.07%. Although this method does not consider matching the fingerprints, the comparison shows the superiority of convolutional neural networks in interacting with fingerprint patterns. This superiority in performance is the result of the ability of neural networks to learn intra- and inter-class variation, so that, more robust decisions can be made by these networks (Michelsanti et al., 2017).

In this paper, a novel method is proposed to measure the similarity between fingerprint images using convolutional neural networks. The proposed method extracts features directly from the pixels' information if the fingerprint image in order to measure the similarity

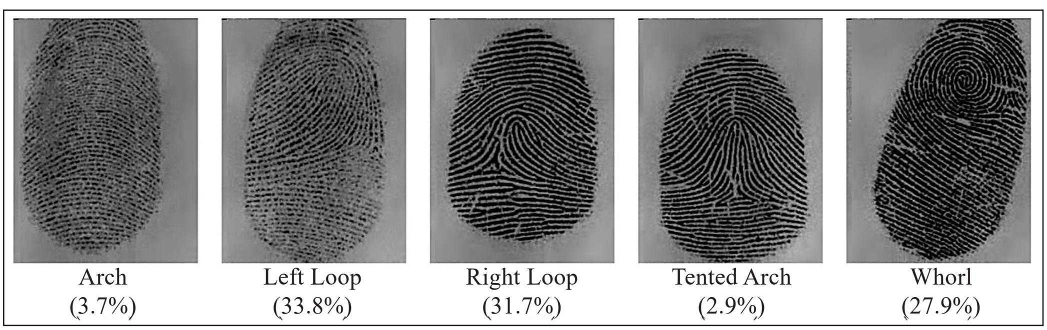


Figure 1. Fingerprints classes defined by Henry, (1905) and their frequencies

between these images. Hence, the features extraction and descriptors generation stages are fused, which eliminates the risk of manipulating these features or descriptors and produce false matches. Moreover, according to the ability of the neural networks to learn the inter- and intra-class variation, the proposed method can produce a better decision, compared to the use of computer-vision techniques like SURF. However, as the proposed method measures the similarity between fingerprint images, instead of classifying them, it is possible to use this method on any datasets without the need to retrain the neural network when the individuals in the dataset change, i.e. the number of neurons in the output layer is constant regardless of the number of individuals in the model fingerprints database.

METHOD

As the proposed method is required to process two fingerprint images and produce a single value that represents the similarity between these two fingerprint images, the input layer of the implemented neural network is required to accept a three-dimensional array while the output layer contains a single neuron. The three-dimensional input contains two fingerprint images, each represented by a two-dimensional array. The similarity measure outputted by the neuron in the output layer is limited in the interval [0,1]. Hence, the activation function used in this neuron was the sigmoid function, which produced values within the required interval, shown in Equation 1. Moreover, according to the significant improvement in the performance of neural networks when the Rectified Linear Unit (ReLU) activation function is used in the neurons of the hidden layers, this activation function is employed in the corresponding neurons(Zhang et al., 2014).

$$\sigma(x) = \frac{1}{1+e^{-x}} \tag{1}$$

$$ReLU(x) = \begin{cases} x & x \geq 0 \\ 0 & x < 0 \end{cases} \tag{2}$$

As shown in Figure 2, the shape of the inputs delivered to the convolutional neural network was 200×200×2, i.e., two images with 200×200 pixels each. This input layer was followed by three convolutional layers, with 32, 16 and 8 filters in each, sequentially, where each filter had a size of (10×10), (7×7) and (3×3), for these three layers. Each convolutional layer was followed by a Max-Pooling layer with the size of 2×2 to emphasize the strong features and maintain accurate positioning. The output of the last Max-Pooling layer was flattened and connected to four hidden fully-connected layers, with 512, 256, 128 and 64 neurons, each. Table 1 describes the details of each layer in the implemented neural network.

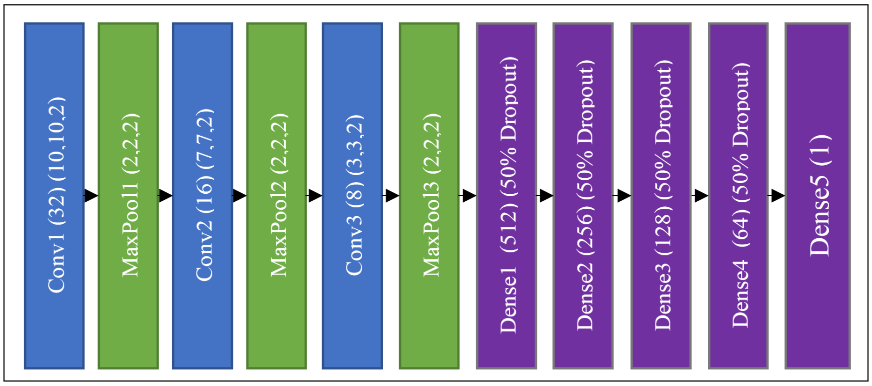


Figure 2. Topology of the convolutional neural network implemented for the proposed methoda

Table 1
Description of the layers in the implemented neural network.

Layer	Input Shape	Neurons (filters)	Activation
Input	200×200×2	-	-
Conv1	200×200×2	32 (10×10×2)	ReLU
MaxPool1	191×191×32	(2×2×2)	-
Conv2	95×95×32	16 (7×7×2)	ReLU
MaxPool2	89×89×16	(2×2×2)	-
Conv3	44×44×16	8 (3×3×2)	ReLU
MaxPool3	(42×42×8)	(2×2×2)	-
Flatten	(21×21×8)	3528	-
Dense1	3528	512	ReLU
Dense2	512	256	ReLU
Dense3	256	128	ReLU
Dense4	128	64	ReLU
Dense5	64	1	Sigmoid

According to the ability of artificial neural networks in recognizing the variation in the inter- and intra-class, the output of neural network is trained to produce the probability of the input fingerprints to be for the same individual, instead of producing an absolute similarity measure as in the use of standard computer-vision techniques, such as SURF. Hence, fingerprint images pairs that belong to the same individual were labeled with one, while pairs of fingerprint images from different individuals were labeled with zero during the training of the neural network. Labels with values of ones indicate 100% confidence that the pairs belong to the same individual, while the zeros indicate 0% confidence that the pair contains fingerprint images of the same individual. Using such approach, the convolutional neural network extracts the knowledge of how to match fingerprints, instead of classifying them, so that, the same trained model can be used with other pairs, that have never been included in the training. As the output of the network is the probability of the fingerprints to belong to the same individuals, new individuals can be recognized by the proposed method by simply including model images of their fingerprints in the pairs inputted to the neural network. Figure 3 shows samples of the inputs and outputs of the convolutional neural network, using the proposed approach.

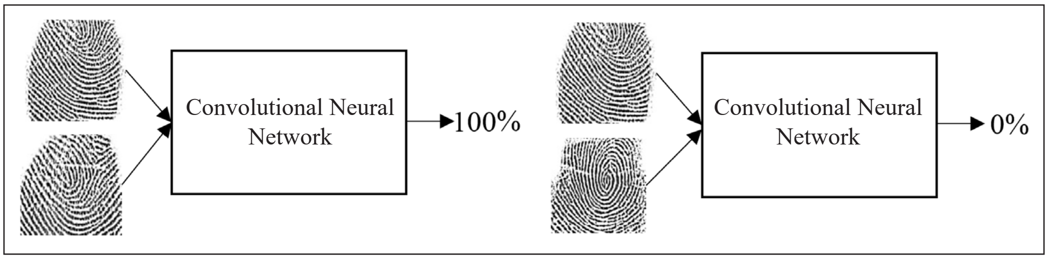


Figure 3. Similarity measurement using the proposed method. Left: fingerprint images of the same individual; Right: fingerprint images for different individuals

To train the neural network for the intended application, triplet loss was used, which was widely used to train neural network for biometric recognition and authentication applications. Per each image in the training dataset, denoted as the anchor image, two additional images were selected from that dataset. One of these images was positive, i.e. was collected from the same individual but was not the same anchor image. The other fingerprint image was the negative, which was collected from any other individual than the one that the anchor image belonged to. Hence, the number of training pars was twice the number of images in the training dataset, as per each image a positive and a negative pair were generated.

RESULTS AND DISCUSSION

In order to evaluate the performance of the proposed method, the model was implemented using Python (Sanner, 1999) programming language with a computer that ran on an Intel® Core™ i7-7700HQ CPU at 2.80GHz frequency and a 16GB of random access memory. The computer also had an Nvidia GTX1080Ti Graphical Processing Unit (GPU) with 4GB of memory, which is used to accelerate the performance of the neural network, implemented using Keras (Chollet, 2015) library, implemented on top of Google's Tensorflow (Abadi et al., 2016) machine learning library. The FVC2002_DB1 (Maio et al., 2002) and FVC2004_DB1 (Maio et al., 2004) datasets were used for the training and evaluation of the proposed method. Five individuals per each dataset, i.e. a total of 10 individuals, were excluded from the training phase and used for the evaluation.

The exclusion of 10 individuals from the training dataset, instead of excluding fingerprint images of individuals that were in the training dataset, was to illustrate the ability of the proposed method to predict the authenticity of images from individuals that were never included in the training. Each image in the training and testing dataset was paired with all the images in the same dataset, including itself. Pairs of the same individual were labeled with one, and the others were labeled with zeros. The neural network was trained for 1000 epochs, before it was evaluated using the testing dataset.

The evaluation of the proposed method was conducted using the confusion matrix shown in Table 2. The threshold value that produced Equal Error Rate was selected, i.e. the False Acceptance Rate (FAR), shown in Equation 3, and False Rejection Rate (FRR), shown in Equation 4, were equal. The value of the threshold that produced the EER was selected based on the Receiver Operating Characteristics (ROC) curve.

Table 2
Confusion matrix of the authentication system

		Predicted	
		Accept	Reject
Actual	Accept	True Acceptance (TA)	False Rejection (FR)
	Reject	False Acceptance (FA)	True Rejection (TR)

$$FAR = \frac{FA}{FA + TR} \quad (3)$$

$$RFRR = \frac{FR}{FR + TA} \quad (4)$$

As each individual in these datasets had 8 fingerprint images, the evaluation dataset contained 6400 pairs, 640 from the same individual and 5760 pairs from different individuals, which represented intrusion attempts. The fingerprint images were resized to 200×200 pixels to match the dimensions of the input layer of the implemented neural network and reduced the complexity of computations. Figure 4 shows the ROC curve of the proposed method for the evaluation dataset. This figure illustrates the ability of artificial

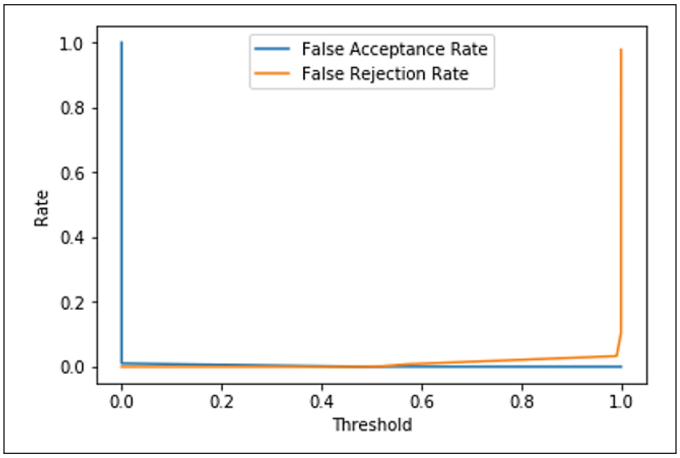


Figure 4. ROC curve of the proposed fingerprint-based authentication method

neural networks to recognize the inter- and intra-class variation, so that, extremely low values are produced per each pair of fingerprint images where each image is collected from a different user. The values produced for fingerprint image pairs collected from the same individual are extremely high, i.e. close to 100%. These values show that the output of the neural network represents the probability of the pair to be for the same individual rather than an absolute similarity measure between the images.

The threshold value that had been able to achieve EER was 0.4754, which had produced confusion matrix shown in Table 3. The accuracy of the authentication decision using the proposed method was 99.97%, with 0.035% FAR and 0% FRR. Although the FAR and FRR are not equal, these are the most similar values that the ROC curve has been able to produce, where selecting different threshold value increases the gap between the values dramatically. Hence the EER of the proposed method is calculated as the average of the FAR and FRR, which is 0.018%, similar to Kumar et al. (2016).

Table 3
Confusion matrix of the authentication decisions of the proposed method at a threshold value of 0.4754

		Predicted	
		Accept	Reject
Actual	Accept	640	0
	Reject	2	5758

Per each individual, the performance measures of the proposed method are shown in Table 4, which shows that the errors occur with a single individual in the entire testing dataset. A comparison with the method proposed by Kumar et al. (2016), which uses SURF-based matching techniques and uses the same datasets for the evaluation, shows that the

Table 4
Evaluation parameters per each individual in the testing dataset

Individual	Evaluation Parameters (%)			
	FAR	FRR	EER	Accuracy
Ind.1	0	0	0	100
Ind.2	0	0	0	100
Ind.3	0	0	0	100
Ind.4	0	0	0	100
Ind.5	0	0	0	100
Ind.6	0	0	0	100
Ind.7	0	0	0	100
Ind.8	0.35	0	0.175	99.69
Ind.9	0	0	0	100
Ind.10	0	0	0	100
Average	0.035	0	0.018	99.97

proposed method has better performance. The method proposed by Kumar et al., (2016) had 99.4% average accuracy with 0.03% average FAR and 0.05% average FRR. Moreover, as the proposed method computes the probability of input fingerprint images to be from the same individual directly from the pixels' information, the vulnerabilities, imposed by extracting features and matching them in different stages, are eliminated.

CONCLUSION

This work proposes a similarity measurement technique for fingerprint images using a convolutional neural network. The use of such approach combines the accuracy of these networks with the flexibility of matching approach, instead of the default classification approach that these networks are usually used for. The results of the performance evaluation experiments illustrate these features, where a matching accuracy of 99.97% is achieved by the neural with 0.035% FAR and 0% FRR. Moreover, the proposed method has been able to outperform the state-of-the-art technique existing in the literature while maintaining high security.

In future work, the application of the same approach is going to be evaluated on different biometric authentication systems, such as face and iris recognition. Such an application can significantly improve the performance of these authentication systems on both security and usability measures.

ACKNOWLEDGEMENT

The authors would like to thank the anonymous reviewers for their constructive comments to improve the quality of this paper.

REFERENCES

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., ... & Kudlur, M. (2016, November 2-4). Tensorflow: A system for large-scale machine learning. In *12th (USENIX) Symposium on Operating Systems Design and Implementation (OSDI' 16)* (pp. 265-283). Savannah, GA, USA.

Alotaibi, A., & Mahmmod, A. (2015, May 1). Enhancing OAuth services security by an authentication service with face recognition. In *2015 Long Island Systems, Applications and Technology* (pp. 1-6). Farmingdale, NY, USA.

Barni, M., Droandi, G., & Lazzeretti, R. (2015). Privacy protection in biometric-based recognition systems: A marriage between cryptography and signal processing. *IEEE Signal Processing Magazine, 32*(5), 66-76.

Chollet, F. (2015). *Keras: Deep Learning for Humans*. GitHub repository. Retrieved November 15, 2018 from https://github.com/fchollet/keras.

Douglas, M., Bailey, K., Leeney, M., & Curran, K. (2018). An overview of steganography techniques applied to the protection of biometric data. *Multimedia Tools and Applications, 77*(13), 17333-17373.

Henry, E. R. (1905). *Classification and uses of Finger Prints*. London, England: HM Stationery Office.

Kumar, R., Chandra, P., & Hanmandlu, M. (2016). A robust fingerprint matching system using orientation features. *Journal of Information Processing Systems, 12*(1), 83-99.

Maio, D., Maltoni, D., Cappelli, R., Wayman, J. L., & Jain, A. K. (2002, August 11-15). FVC2002: Second fingerprint verification competition. In *Object recognition supported by user interaction for service robots* (Vol. 3, pp. 811-814). Quebec City, Quebec, Canada.

Maio, D., Maltoni, D., Cappelli, R., Wayman, J. L., & Jain, A. K. (2004, July 15-17). FVC2004: Third fingerprint verification competition. In *International Conference on Biometric Authentication* (pp. 1-7). Hong Kong, China.

McAteer, I., Ibrahim, A., Zheng, G., Yang, W., & Valli, C. J. T. (2019). Integration of biometrics and steganography: A comprehensive review. *Technologies, 7*(2), 34-55.

Michelsanti, D., Guichi, Y., Ene, A. D., Stef, R., Nasrollahi, K., & Moeslund, T. B. (2017, February 27 - March 1). Fast fingerprint classification with deep neural network. In *Proceedings of the 12th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2017)* (pp. 202-209). Porto, Portugal.

Nagatomo, M., Kita, Y., Aburada, K., Okazaki, N., & Park, M. (2018). Implementation and user testing of personal authentication having shoulder surfing resistance with mouse operations. *IEICE Communications Express, 7*(3), 77-82.

Najih, A., Al-Haddad, S. A. R., Ramli, A. R., Hashim, S. J., & Nematollahi, M. A. (2016). An overview of multimodal biometric approaches based on digital image watermarking. *Research Journal of Applied Sciences, Engineering and Technology, 13*(6), 481-494.

Orrù, G., Casula, R., Tuveri, P., Bazzoni, C., Dessalvi, G., Micheletto, M., ... & Marcialis, G. L. (2019, In Press). LivDet in Action-Fingerprint Liveness Detection Competition 2019. *arXiv preprint arXiv:1905.00639*.

Peralta, D., Triguero, I., García, S., Saeys, Y., Benitez, J. M., & Herrera, F. (2018). On the use of convolutional neural networks for robust classification of multiple fingerprint captures. *International Journal of Intelligent Systems, 33*(1), 213-230.

Ratha, N. K., Connell, J. H., & Bolle, R. M. (2001). Enhancing security and privacy in biometrics-based authentication systems. *IBM Systems Journal, 40*(3), 614-634.

Sanner, M. F. (1999). Python: a programming language for software integration and development. *Journal of Molecular Graphics and Modelling, 17*(1), 57-61.

Sinha, H., & Ajmera, P. K. (2019). Upgrading security and protection in ear biometrics. *IET Biometrics, 8*(4), 259-266.

Sun, H. M., Chen, S. T., Yeh, J. H., & Cheng, C. Y. (2018). A shoulder surfing resistant graphical authentication system. *IEEE Transactions on Dependable and Secure Computing, 15*(2), 180-193.

Wang, K., Jiang, J., Cao, Y., Xing, X., & Zhang, R. (2016, November 5-7). Preprocessing Algorithm Research of Touchless Fingerprint Feature Extraction and Matching. In *Chinese Conference on Pattern Recognition* (pp. 436-450). Chengdu, China.

Zhang, S., Bao, Y., Zhou, P., Jiang, H., & Dai, L. (2014, May 4-9). Improving deep neural networks for LVCSR using dropout and shrinking structure. In *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 6849-6853). Florence, Italy.

Zou, Y., Zhao, M., Zhou, Z., Lin, J., Li, M., & Wu, K. (2018). BiLock: User Authentication via Dental Occlusion Biometrics. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2*(3), 1-20.

Pertanika J. Sci. & Technol. 27 (4): 1735 - 1758 (2019)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Improved Salient Object Detection via Boundary Components Affinity

Nur Zulaikhah Nadzri[1,2]*, Mohammad Hamiruce Marhaban[1], Siti Anom Ahmad[1], Asnor Juraiza Ishak[1] and Zalhan Mohd Zin[2]

[1]*Department of Electrical and Electronic Engineering, Faculty of Engineering, Universiti Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia*

[2]*Industrial Automation Section, Universiti Kuala Lumpur Malaysia France Institute, Section 14, 43560, Bandar Baru Bangi, Selangor, Malaysia*

ABSTRACT

Referring to the existing model that considers the image boundary as the image background, the model is still not able to produce an optimum detection. This paper is introducing the combination features at the boundary known as boundary components affinity that is capable to produce an optimum measure on the image background. It consists of contrast, spatial location, force interaction and boundary ratio that contribute to a novel boundary connectivity measure. The integrated features are capable to produce clearer background with minimum unwanted foreground patches compared to the ground truth. The extracted boundary features are integrated as the boundary components affinity. These features were used for measuring the image background through its boundary connectivity to obtain the final salient object detection. Using the verified datasets, the performance of the proposed model was measured and compared with the 4 state-of-art models. In addition, the model performance was tested on the close contrast images. The detection performance was compared and analysed based on the precision, recall, true positive rate, false positive rate, F Measure and Mean Absolute Error (MAE). The model had successfully reduced the MAE by maximum of 9.4%.

Keyword: Boundary connectivity, boundary ratio, force interaction

ARTICLE INFO

Article history:
Received: 18 January 2019
Accepted: 13 May 2019
Published: 21 October 2019

E-mail addresses:
eka.nadzri@gmail.com (Nur Zulaikhah Nadzri)
mhm@upm.edu.my (Mohammad Hamiruce Marhaban)
sanom@upm.edu.my (Siti Anom Ahmad)
asnorji@upm.edu.my (Asnor Juraiza Ishak)
zalhan@unikl.edu.my (Zalhan Mohd Zin)
* Corresponding author

INTRODUCTION

Salient detection is the ability to detect the most prominent object on a particular scene or region. This is adapted from the ability of human eyes to distinguish distinctive

ISSN: 0128-7680
e-ISSN: 2231-8526

(salient) objects on the visual field. Out of large inputs entering our eyes, most of it are filtered and the ones left are useful for cognitive purpose. The factor that influences visual saliency lies in two mechanisms, which are bottom-up attention and top-down attention. The bottom-up mechanism is fast and stimulus-driven attention in which the visual saliency is based on the low-level features including color, intensity, orientation, texture, and motion. On the other hand, the top-down mechanism is slower as it is goal-driven which referring to internal guidance of attention based on prior knowledge (Katsuki & Constantinidis, 2014) . It is a voluntary allocation of attention to certain features, objects, or regions in space (Pinto et al., 2013)

There have been a large number of studies in salient object detection in recent years. This area has attracted researchers in which hundreds of computational models were proposed to obtain the most real-time and optimum result on the detection.

The salient object detection model has been applied in many applications such as object detection and segmentation (Liu et al., 2008; Xiuli et al., 2017; Zhang et al., 2006), image retargeting (Chou & Su, 2016; Pritch et al., 2013), image compression (Guo & Zhang, 2010; Itti, 2004; Srivastava et al., 2016) and image quality assessment (Xiao & Yeh, 2017). The models have been successfully applied in many multi-discipline areas including multimedia (Li et al., 2017; Luz et al., 2017), medical (Woodbridge et al., 2011; Ahn et al., 2017), remote sensing (Li et al., 2016; Zhang et al., 2016) and robotics (Jiang et al., 2015; Liu et al., 2013).

The earliest salient detection model was proposed by Itti et al. (2001) which was based on the selective attention mapping by Koch and Ullman (1985) together with the basis of other models of Baluja (1997) and Milanese et al. (1995). It is also related to the physiological theories of visual attention called the Feature Integration Theory by Treisman and Galade (1980) that explains the human visual strategies. The theory suggests that the attention onto an object involved in separate process where the early perceived stimulus will put all the features in parallel as a pre-attentive stage and the individual features are combined to select the focus location. The model has become hit in a way that it was used as a basis for multi cross discipline in cognitive psychology, neuroscience, and computer vision.

In recent years, multiple approaches have been used as the basis of hypothesis on the salient detection model in which the base theorem is always being referred to the term 'prior' of the presented models. To list a few, which are contrast prior to either local (Achanta et al., 2008; Itti et al., 2001) or global (Cheng et al., 2011; Perazzi & Kr, 2012; Zhang et al., 2016), edge/shape prior (Jiang et al., 2011; Yang et al., 2017), texture prior (Hu et al., 2016; Zhang et al., 2017), background prior (Ahn et al., 2017; Jing et al., 2014; Wei et al., 2012), or foreground prior (Wang et al., 2015; Zhou et al., 2017). The contrast prior has been the most applied assumption in which the salient object is highlighted whenever the contrast between the object and the surroundings are high. Besides, boundary prior models (Luo et al., 2016; Manke & Jalal, 2016; Niu, 2018; Tang et al., 2018; Wu et al., 2013) have proven good results in salient detection where the image boundary is being assumed as the image background.

The earlier model that has exploited the boundary prior is by Wei et al. (2012) . The model is related to the assumption that associates image boundary and their connectivity within the image patches. The saliency detection is improved with the integration of geodesic saliency that measures the shortest path of the image patches to the background patch. Motivated by the model, Zhu et al. (2014) had come out with saliency detection that relied on the measure of the image patches as background only when the region it belonged to was strongly connected to the image boundary. The model by Wu et al (2013) uses the *Lo* smoothing filter and the Principal Component Analysis to make categorization of the salient object, and at the same time produces the boundary information for the background merging and boundary scoring stages. On adapting the boundary prior to salient detection model, Manke et al. (2016) used the Poisson distribution to highlight the salient object by computing the difference of pixel intensity and mean of boundary pixel of an image. Considering other perspectives within the boundary prior to the assumption, the models by Luo et al., (2016); and Tang et al., (2018) are focusing on the probability that the salient object could be located at the image boundary.

Among those models, the model by Zhu et al (2014) has been listed as the top 6 models for the salient detection supporting along a detailed study comparison by Borji et al. (2015). With proven computation for the saliency detection, the detection of the model still falsely shows some image patches that should be assigned as the image background. It is due to their boundary connectivity computation that measures the contrast difference towards the whole image boundary and not considering the weight or ratio of the patches towards the individual side of the boundary. Thus, the large contrast difference of patches will be considered as the image foreground even though they are highly connected to the image boundary. As a result, any high connectivity of patches at the boundary with higher contrast difference will be highlighted as the salient object. Figure 1 shows the saliency map by Zhu et al. (2014) named as RBD where other high contrast patches were pop-out as the salient object.

This paper presents a new method for the saliency detection based on the image boundary prior. The main contribution is the combination features extracted that consist of contrast, spatial location, force interaction and boundary ratio that is assigned as the Boundary Components Affinity (BCA). Apart from only taking into account the patches' contrast and location difference using the Euclidean distance, considering the patch force interactions is able to strengthen the patch differences in grouping them according to the object they belong (foreground or background). On the other hand, the boundary ratio is

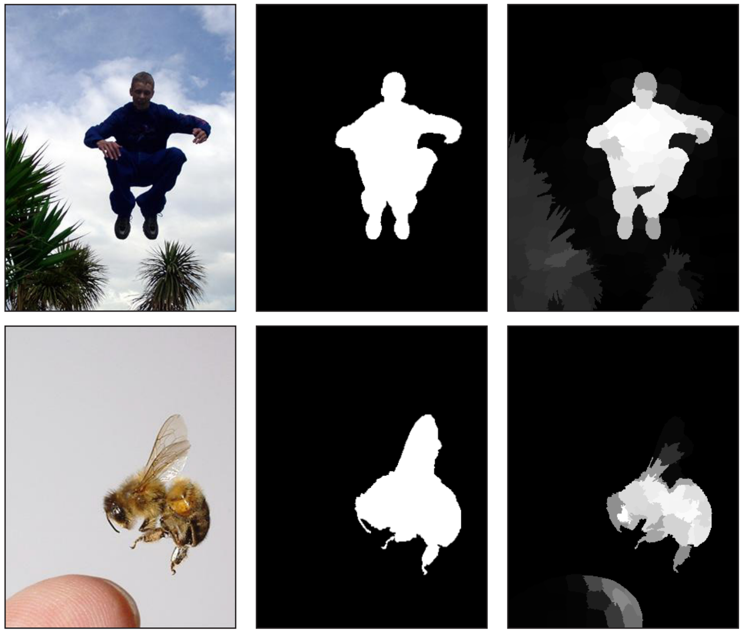


Figure 1. (From left to right) Original image, ground truth and saliency map of RBD (Zhu et al., 2014)

used to identify the ratio of a patch connected to which image boundary. These features are integrated as the BCA and act as a measure of the boundary connectivity. This measure is able to assign patches as foreground or background of the image. Higher boundary connectivity means that the patch is highly connected to the image boundary and hence it would be assigned as the image background. With these combination features, the proposed saliency measure is computed from novel integrated boundary components in which this method is able to highlight the salient object by producing the optimum background measure.

THE PROPOSED FRAMEWORK OF BCA

The following section discusses the detail of the proposed salient detection algorithm. The core of this saliency detection arises from the background measurement. The measurement is made by integrating the contrast, spatial location, force interaction and boundary ratio of the patches to be put together as the boundary components and hence measure the connectivity of the patches towards the image boundary. The spatial force measurement is able to highlight the salient object by obtaining the force of each patch related to its spatial contrast and distance. Thus, the high force of the patches is grouped based on the object it belongs to. The boundary ratio is used to strengthen the computation of patches-boundary connectivity. The detail about the algorithm will be discussed in the next section. The summary of the stated algorithm is depicted as in Figure 2. It is inspired by the RBD model (Zhu et al., 2014) where the extracted BCA are integrated for computing the boundary connectivity as they are highlighted in red dotted square which illustrates the additional steps that have been added in the proposed BCA in comparison with the RBD model. The details of the proposed algorithm are discussed in the following subsections.

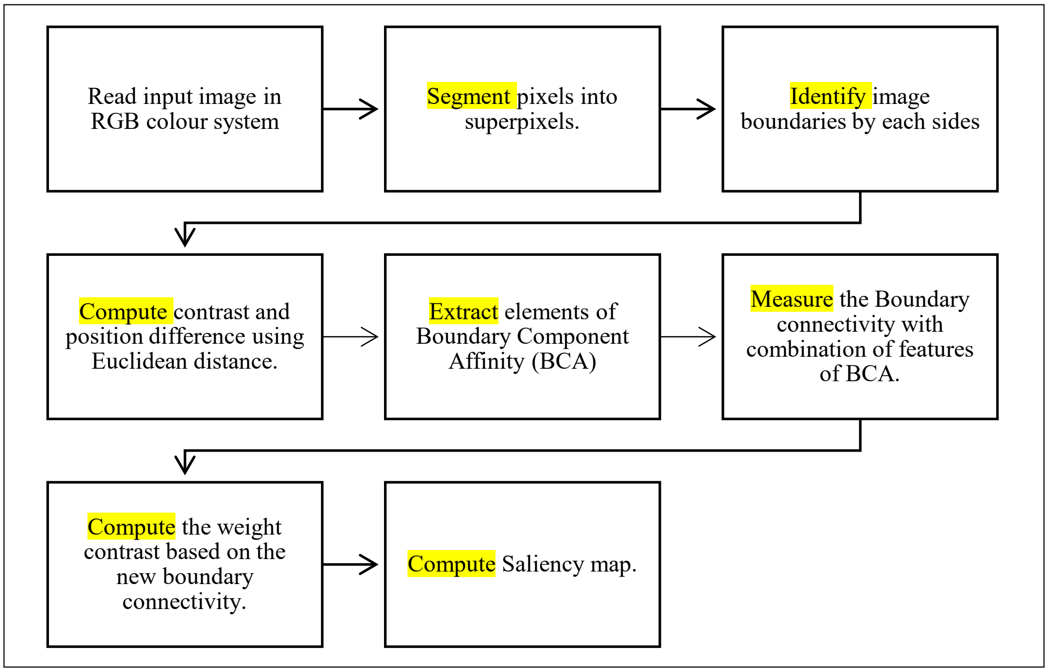


Figure 2. The proposed model block diagram

Pre-processing

Having the fastest detection time is the vital criteria for any salient detection algorithm as processing a large number of pixels, i.e. 300 x 400 pixels, may result in a higher computational time. Thus, the image is segmented using the Simple Linear Iterative Clustering (SLIC) technique (Achanta et al., 2012) where clustering the pixels is 5-dimensional including color and plane space to obtain a compact superpixel. The unique distance measure produces compactness and regularity in the superpixel shapes and thus resulting an efficient superpixel segmentation. With a good resulting segmentation, the images will be in accurate group of patches with reduce number of component to be processed as the saliency measurement rather that individual pixels. Figure 3 shows the result of SLIC superpixel segmentation.

Figure 3. (From left to right) Original image and superpixel segmentation using SLIC

Pixel Force Feature

The concept of spatial force is taken from the concept of pixel-force field (Hurley et al., 1999). The mathematical modelling of object force interaction in physics is adopted for the pixel force transformation. This feature has been demonstrated to benefit numbers of application such as segmentation (Bucha et al., 2007), extraction and recognition (Bucha et al., 2006; Hurley et al., 1999), map vectorization (Bucha et al., 2007), and image registration (Ghayoor, 2010). Each pixel in an image is considered as a single particle that contains a specific scalar and vector value towards other pixels(Hurley et al., 1999) and it can be visualized as in Figure 4. In transforming image into the force field, each pixel is presumed to produce symmetrical force field, $\mathbf{F}_i(\mathbf{r})$ towards other pixel of $P(\mathbf{r}_i)$ where $\mathbf{r}$ is the location vector. It can be defined mathematically as Eqation [1],

$$\mathbf{F}_\mathrm{i}(\mathbf{r}) = P(\mathbf{r}_\mathrm{i})\frac{\mathbf{r}_\mathrm{i} - \mathbf{r}}{|\mathbf{r}_\mathrm{i} - \mathbf{r}|^3} \qquad [1]$$

Generally, a single pixel p_i in an image can be presented as their two dimensional spatial location vectors and the color appearance q in n dimensional vector, $\mathbf{Q_p} = (q1, \ldots, qn)$. The definition of the pixel interaction force, $\mathbf{F}_{i,j}$ can be referred to Bucha et al (2007). The force, $\mathbf{F}_{i,j}$ between two pixels p_i and p_j is defined by the pixel spatial relation $D(\mathbf{X_p}(i), \mathbf{X_p}(j))$

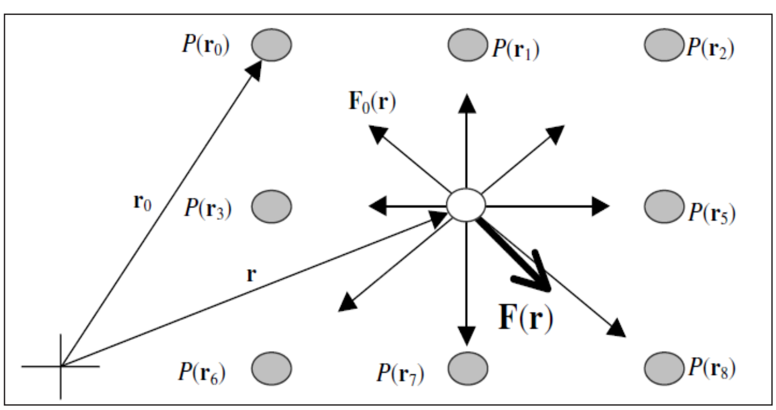


Figure 4. Illustration of force field geometry (Hurley et al., 1999)

and their color dissimilarities $C(\mathbf{Q_p}(i), \mathbf{Q_p}(j))$. Thus, the pixel interaction force between p_i and p_j can be presented as Equation2.

$$\mathbf{F}(p_i, p_j) = D(\mathbf{X_p}(i), \mathbf{X_p}(j)) \cdot C(\mathbf{Q_p}(i), \mathbf{Q_p}(j)) \cdot \mathbf{r}_{i,j} \quad [2]$$

The two dimensional vector, $\mathbf{r}_{i,j}$, is the direction of the pixel interaction force from the coordinate location of pixel, p_i, to the coordinate location of pixel, p_j.

Image Patch Force Interaction

In applying this pixel force field interaction into the model, the pixels are replaced with the superpixel patches which are obtained from the superpixel segmentation earlier. Note that the patch force interaction is presented as $\mathbf{F}(P_i, P_j)$ where P_i is a particular patch of the superpixel. As a single patch consists of a number of pixels, the location can be presented as the mean spatial location $\bar{\mathbf{X}}_\mathbf{P}(i) = (\bar{x}_a, \bar{y}_a)$ where $a \in i$. The component a is the set of pixels that are bounded in the patch i.

To imitate the human vision, the CIE Lab color system is used for the color vector where the pixel appearance, $\mathbf{Q}_p$ is presented such that,

$$\mathbf{Q_p} = (q_L, q_a, q_b) \quad [3]$$

Where q_L, q_a and q_b are pixel appearances on the L, a and b channel.

However, the vector needs to be presented as a vector for the superpixel patch $\bar{\mathbf{Q}}_\mathbf{P} = (\bar{q}_L, \bar{q}_a, \bar{q}_b)$ as the mean of CIE Lab color of a pixel set a of the patch i originates from the RGB input pixels.

The scalar force interaction between the patches P_i and P_j is obtained by substituting the relevant values in equation (2) and presented as in Equation [4].

$$F_{ij} = D(\bar{X}_P(i), \bar{X}_P(j)) \cdot K(\bar{Q}_P(i), \bar{Q}_P(j)) \quad [4]$$

Both elements $D(\bar{X}_P(i), \bar{X}_P(j))$ and $K(\bar{Q}_P(i), \bar{Q}_P(j))$ are the Euclidean distance of the spatial coordinate and CIE Lab difference between patches respectively which are based on the equation introduced by Bucha et al. (2007). They can be written as Equations [5], [6] and [7].

$$D(\bar{X}_P(i), \bar{X}_P(j)) = \frac{1}{d(\bar{X}_P(i), \bar{X}_P(j))} = \frac{1}{\sqrt{\sum_{i=1}^{n}\left(|\bar{x}_i - \bar{x}_j|\right)^2 + \left(|\bar{y}_i - \bar{y}_j|\right)^2}} \quad [5]$$

$$K(\bar{Q}_P(i), \bar{Q}_P(j)) d(\bar{Q}_P(i), \bar{Q}_P(j)) = \sqrt{\sum_{i=1}^{n}\left(\bar{q}_{L_i} - \bar{q}_{L_j}\right)^2 + \left(\bar{q}_{a_i} - \bar{q}_{a_j}\right)^2 + \left(\bar{q}_{b_i} - \bar{q}_{b_j}\right)^2} \quad [6]$$

Combining both equations [5] and [6] into [4], the final force interaction of each patch, $\mathbf{F}_i(p)$ can be presented by,

$$\mathbf{F}_i(P) = \sum_{i=1}^{n} \frac{\sqrt{\sum_{i=1}^{n} \left(\bar{q}_{L_i} - \bar{q}_{L_j}\right)^2 + \left(\bar{q}_{a_i} - \bar{q}_{a_j}\right)^2 + \left(\bar{q}_{b_i} - \bar{q}_{b_j}\right)^2}}{\sqrt{\sum_{i=1}^{n} \left(\left|\bar{x}_i - \bar{x}_j\right|\right)^2 + \left(\left|\bar{y}_i - \bar{y}_j\right|\right)^2}} \cdot \mathbf{r}_{ij} \qquad [7]$$

The calculated force interaction is drastically able to group the similar patch and hence it is able to contribute to the coarse object detection. It is because the force interaction is combining the color feature and spatial feature in one single value. This feature is not integrated in the previous models whereby those models only use the Euclidian distance on both color feature and spatial coordinate separately. The equation in [7] is used in the proposed model by ignoring the vector force direction, $\mathbf{r}_{i,j}$ since this is not an important element to be considered for computation.

Boundary Ratio

In computing saliency, this model considers how near the location of a patch to a specific boundary. The concept is illustrated as in Figure 3. The boundary ratio is the ratio of the patch center (red dot of each patch) to the nearest boundary either at the top, left, right or down boundary. This boundary ratio is used to improve the measure of the boundary connectivity.

The term boundary connectivity has been introduced by Zhu et al. (2014). The boundary connectivity, *b* of a patch, *P* can be measured by taking the ratio of pixels, *p* of a patch , *P* at the image boundary, *m* to the square root of the pixels at patch region, *R* (Equation [8]).

$$b(P) = \frac{|\{p|P \ \in m\}|}{\sqrt{|\{p|P \ \in R\}|}} \qquad [8]$$

From the defined equation, we can see that the measure of boundary connectivity is taking all the four side boundaries into account and this will result in average measure on the boundary connectivity.

To have a more precise computation, the measurement on boundary connectivity is integrated with the boundary ratio to strengthen the patches that are highly connected to the boundary. Here, the model is taking the center of each patch and obtain its nearest distance to a specific boundary. The boundary ratio, b_r of a patch, *P* is defined as the ratio of pixels, *p* of patch *P* towards a specific boundary, *m* to its center position with respect to a boundary side, *m*, whether it referred to top, down, left or right side.

Figure 5 is the illustration on boundary ratio computation. Consider that each color represents a single patch. By locating the mean location of each patch, the distance from the patch center to the nearest boundary is measured. As the patch has the lowest contrast towards the boundary and it is near the patch to any boundary side, the larger the ratio. Therefore, any low contrast patch that is close to the boundary has higher probability to be accounted as the image background. This value will be used to strengthen the boundary connectivity of each patch from the equation [8]. This is the missing consideration in the previous model (Zhu et al., 2014) where the boundary connectivity is measured towards all boundaries. As a result, any small high contrast patches will appear as salient region as in Figure 1.

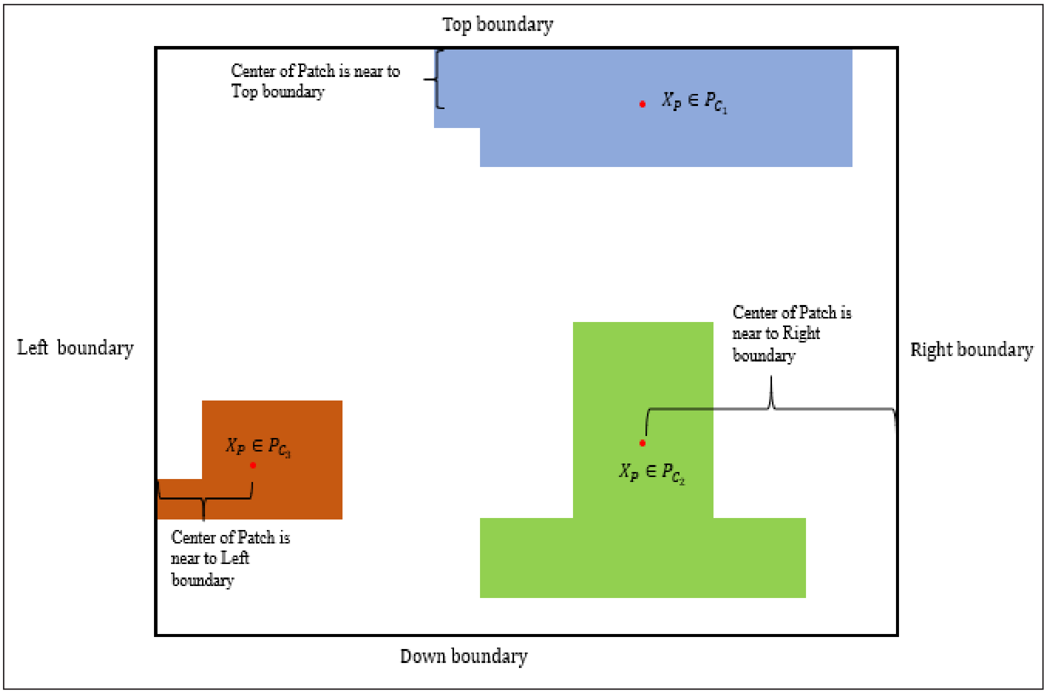


Figure 5. The boundary ratio of each patch towards its nearest boundary

Instead of taking all boundaries in the computation, the ratio of patches towards each boundary, b_r is computed at different boundary side, *m* which refers to top, down, left and right boundaries, respectively. The boundary ratio is calculated by taking the summation of mean appearance difference of patches, *P* and P_i towards a specific boundary *m*, over the distance of the patch mean location towards the boundary which is given by,

$$b_r(P) = max \sum_{i=1}^{N} \frac{\bar{q}(P, P_i) \in m}{(X_P | P_c \in m)} \qquad [9]$$

Then, the maximum of all ratio is obtained to know that the patch is near to which boundary. By referring to the concept of boundary connectivity from (Zhu et al., 2014), the patch on boundary length and the patch area are computed from the geodesic distance between two superpixels $g(p_i, p_j)$, where q is the appearance difference between the neighboring patches. It can be defined as,

$$g(p_i, p_j), = \min_{pi=p,p2,\dots pn=q} \sum_{i=1}^{n-1} q(p_i, p_{i+1}) \quad [10]$$

The geodesic distance is a method that is used to compute the shortest path between the patch on the graph (Wei et al., 2012). Thus the patch area, *a* and patch length at the boundary, *l* are computed by,

$$a(P) = \sum_{i=1}^{n} e^{\left(\frac{-g^2(p,p_i)}{2\tau^2}\right)} = H(P, P_i\Big) \quad [11]$$

$$l(P) = \sum_{i=1}^{N} H(P, P_i). \delta(P_i \in m) \quad [12]$$

The value of τ is set to 10 as being experimented by Zhu et al (2014) model. $H(P,P_i)$ is the spanning area of the superpixel patch. The length at the boundary, $l(P)$ is defined from the spanning area with respect to the image boundary, *m*. The constant δ is set to 1 for the patch superpixel at the boundary and 0 otherwise. The new boundary connectivity, *b(P)* are the multiplication of the patch length, $l_m(P)$ with the patch ratio, $b_r(P)$, the patch force interaction, $F_m(P)$ and the inverse square root of patch area, *a(P)*, where *m* is the measurement towards the image boundary. The patch force interaction, $F_m(P)$ *is* taken from the equation in [7]. It is given by,

$$b(P) = \frac{l_m(P)}{\sqrt{a(P)}} \cdot b_r(P) \cdot F_m(P) \quad [13]$$

This is the major contribution of this model where the terms $l_m(P)$, $b_r(P)$, $F_m(P)$, and *a(P)* are the elements of BCA where the component contrast and spatial location are considered explicitly in the geodesic distance computation in equation [10]. The combination in BCA produces a new definition for boundary connectivity that is able to diminish the high contrast non-salient patch that is highly connected to the boundary since the measurement towards the boundary has been strengthened by combination of the features in BCA.

Background Measurement

The core of the background measure relies on the contrast computation where a reliable measurement will contribute to precise salient detection. Many previous models use the region contrast with respect to its neighbor that acts as saliency cue (Cheng et al., 2011; Perazzi & Kr, 2012; Yan et al., 2013). Generally, the weight contrast, *u(P)* from (Zhu et al., 2014) is the summation of contrast difference, *q* between patches *p* and p_i of the superpixel, multiplied with the distance weight, *w* on the superpixel patch. It can be written as,

$$u(P) = \sum_{i=1}^{N} q(P, P_i) \cdot w(P, P_i) \qquad [14]$$

With $\beta = 0.25$ and $z(P,P_i)$ as the spatial distance between patch *P* and $P_{i,}$ the distance weight, *w* is obtained by,

$$w(P, P_i) = \exp\left(\frac{-z(P, P_i)}{2\beta^2}\right) . z(P, P_i) \qquad [15]$$

Zhu et al. (2014) has extended the background weight contrast, d_i which includes the boundary connectivity, *b(P)* as in equation [9] and μ is set to 1. Thus it can be defined as,

$$d_i(P) = 1 - \exp\left(\frac{-b^2(P)}{2\mu^2}\right) \qquad [16]$$

With this new definition, the re-defined weight contrast of each patch *u(P)* based on equation [16] is,

$$u(P) = \sum_{i=1}^{N} q(P, P_i) w(P, P_i) d_i(P) \qquad [17]$$

To have an optimum measure on the background contrast, the new boundary connectivity measure, *b(P)* in equation [16] is used. Thus, the new optimum background contrast can be defined as,

$$d_i(P) = 1 - \exp\left(\frac{-b^2(P)}{2\sigma_b^2}\right) . z(P, P_i) \qquad [18]$$

Saliency Map Computation

The optimum saliency map computation, *M* measure is adopted from Zhu et al., (2014) model where they combined the terms background, d_i foreground, v_i and smoothness, t_{ij} in a single computation as,

$$M = \sum_{i=1}^{N} d_i(P)s_i^2 + \sum_{i=1}^{N} v_i(P)(s_i - 1)^2 + \sum_{i=1}^{N} t_{ij}\left(s_i - s_j\right)^2 \qquad [19]$$

The term foreground, v_i is computed by subtracting 1 from the background, d_i. The combination of the three terms are able to put the superpixel of large background probability to the small value of s_i (close to 0), and the patch with high foreground probability to the large value of s_i (close to 1). In addition, the smoothness term promotes the enhancement of the saliency values. The smoothness term t_{ij} can be obtained by $w_{ij} = \exp\left(\frac{-d_c(p_i,p_j)}{2\sigma_{clr}^2}\right) + \mu$. The constant μ is set to 0.1 that is used to regularize the cluttered image regions.

EXPERIMENTS

The evaluation of the proposed model was tested on the ASD (Achanta et al., 2009) and ECSSD (Li et al., 2013) datasets. Both datasets consisted of 1000 images with the manually segmented ground truth images for the salient object detection model evaluation. The images in ASD datasets were considered simple containing single salient object with clean background while the ECSSD datasets contained semantically meaningful images and were quite complex.

Using these verified datasets, the performance of the proposed model are measured and compared with 4 state-of-the-art salient detection models which are, *Robust Background Detection*, RBD (Zhu et al., 2014), *Saliency Filter*, SF (Perazzi & Kr, 2012), *Geodesic Saliency*, GS (Wei et al., 2012) and *Manifold Ranking*, MR (Yang et al., 2013). Apart from that, the model performance was also tested on the close contrast images that were manually selected from the ASD dataset. This type of image was tested to evaluate the model detection performance as this type of image had become part of unsolved issues in the study of salient detection (Borji et al., 2014).

Out of all 4 models, the RBD, is listed (Borji et al., 2015) to be among the top models for salient object detection based on exhaustive comparison study done. In evaluating the proposed model, the precision-recall (PR), Receiver-Operating Characteristic (ROC), F-Measure and Mean Absolute Error (MAE) were used, where these measurements are universally-agreed and the standard measurement that can be used for evaluating the salient object detection model (Borji et al., 2015).

Models Comparison on Detection

The precision, γ is the ratio of salient pixels accurately detected and allocated as,

$$\gamma = \frac{B \cap G}{B} \qquad [20]$$

On the other hand, recall is the ratio, η of detected salient pixel over its ground truth as,

$$\eta = \frac{B \cap G}{G} \qquad [21]$$

The detail of the measurement can be defined as in the equations [20] and [21], where B is the saliency detected converted in binary mask, based on the saliency image map M, and G is the image ground truth. On the other hand, the Receiver Operating Characteristic (ROC) curve reports the relationship of True Positive Rate (TPR) , ρ and False Positive Rate (FPR), φ of the saliency map S_b and the ground truth, G. They can be finely written as,

$$\rho = \frac{|S_b \cap G|}{|G|} \qquad [22]$$

$$\varphi = \frac{|S_b \cap G|}{|S_b \cap G| + |\overline{S_b} \cap \bar{G}|} \qquad [23]$$

For simplification, comparison is made based on the saliency map S_b that is normalized into [0,255] range. The curve is plotted by taking the average of computed precision and recall on the dataset. The F-measure, F_a and mean Absolute Error (MAE), ε are defined as,

$$F_a = \frac{(1+\alpha)\rho \ \times \eta}{\alpha \times \rho + \eta} \qquad [24]$$

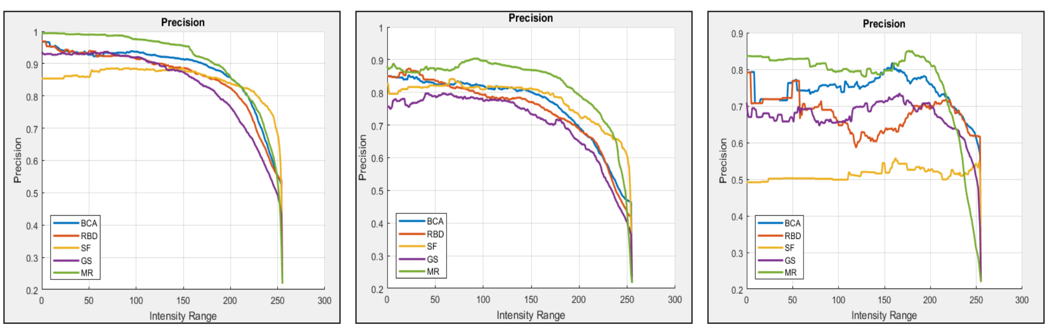
$$\varepsilon = \frac{1}{W \times H} \sum_{x=1}^{W} \sum_{y=1}^{H} \left\| S_b(x,y) - G(x-y) \right\| \qquad [25]$$

Where W and H is the width and height of the image respectively. The constant , α for F_a measurement is set to 0.3 to raise the importance of precision as suggested by Achanta et al. (2009).

Detection Evaluation on Datasets

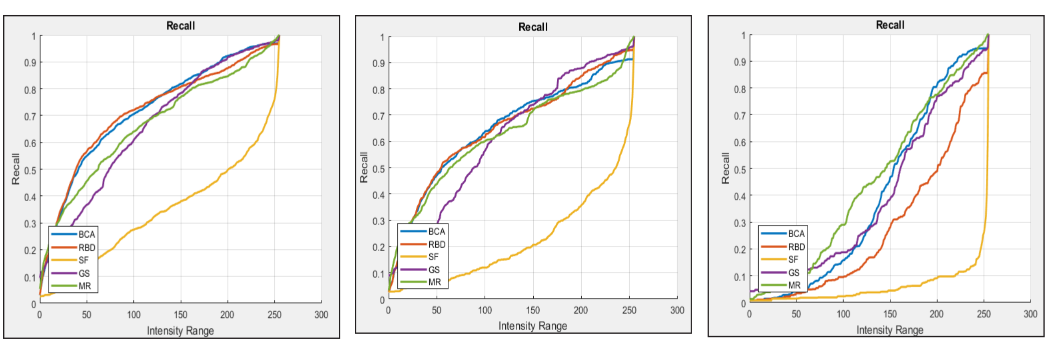
All performance comparisons were put on 3 tests using the ASD and ECSSD datasets as well as the close contrast images collected from the ASD dataset. The performances were compared side by side based on precision and recall with intensity range, precision over recall, and ROC curves. Histograms were used to compare the performances on MAE, precision, recall and F-measure. Samples of tested images, ground truth and saliency were compared as well. In each result comparison, the proposed model is labeled as BCA as the other 4 state-of-the-art models are labeled as RBD, SF, GS and MR. The details of the results analysis are discussed in the next section.

Precision and Recall over the Intensity Range. In these comparisons, the obtained saliency maps are normalized to [0,255] to generate the binary masks. From the curves in Figure 6, it can be observed that the proposed model has high precision compared to other models with MR as the highest. However, the model's precision is low on ECSSD dataset due to its difficult images. The high precision indicates that the saliency map obtained had lower fall-out rate. The curves obtained for close contrast images were quite uneven since there were small number of those types of images in the ASD data set. Referring to the recall curve over the intensity range in Figure 7, the proposed model has the highest and consistent recall value compared to MR on ASD and ECSSD datasets. It indicates that the saliency map obtained from BCA has less miss rate. It can also be observed that the model SF produces an abnormality curve compared to others. It is because its global contrast assumption would result in noisy segmentation as the image has low contrast due to lighting or similar foreground background appearance. Thus, the resulted saliency map would be in high false negative pixels. The min-cut segmentation could be the solution of the issue as discussed in their research paper (Perazzi & Kr, 2012).



(a) (b) (c)

Figure 6. Precision over intensity range on the BCA comparison with other state-of-the-art model using: (a) ASD dataset; (b) ECSSD dataset; and (c) close contrast images collected from ASD dataset.



(a) (b) (c)

Figure 7. Recall over intensity range on the BCA comparison with other state-of-the-art model using: (a) ASD dataset; (b) ECSSD dataset; and (c) close contrast images collected from ASD dataset.

Precision over Recall Curve. Comparing both precision and recall of BCA with other state-of-the-art models in Figure 8, BCA had the second highest precision after MR on the 3 testing datasets. However, BCA had the highest recall when being tested on ASD and close contrast images. The higher and closer the curve to the right, the lesser the miss rate and false alarm of the obtained saliency map. It indicates that the model is able to produce correct salient detection region with lesser false salient region.

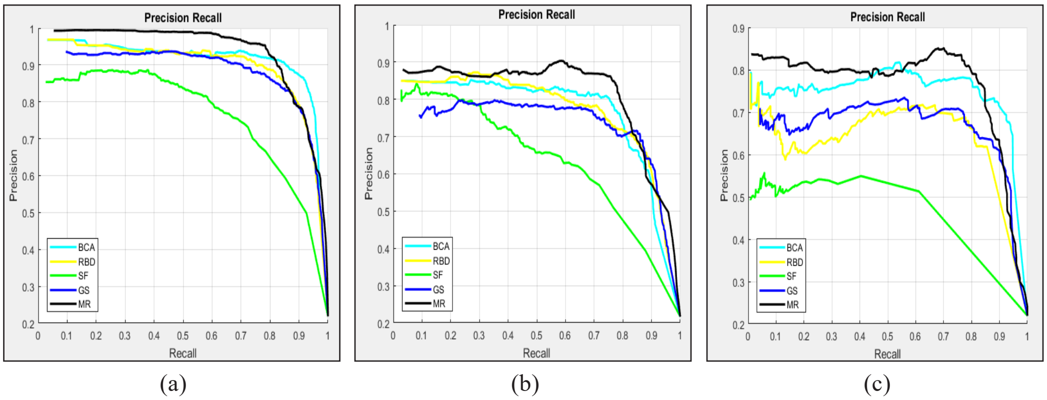


Figure 8. PR curve on the BCA comparison with other 4 state-of-the-art models using: (a) ASD dataset; and (b) ECSSD dataset; and (c) close contrast images collected from ASD dataset.

ROC Curve. In Figure 9, the proposed model BCA had the highest True Positive Rate when being tested with the ASD and close contrast images. It indicates that the salient detection obtained from the model had the highest rate when being compared to its ground truth. The lower curve indicates that the saliency map obtained consisted of many fall-out rates where many patches were falsely detected as the salient region.

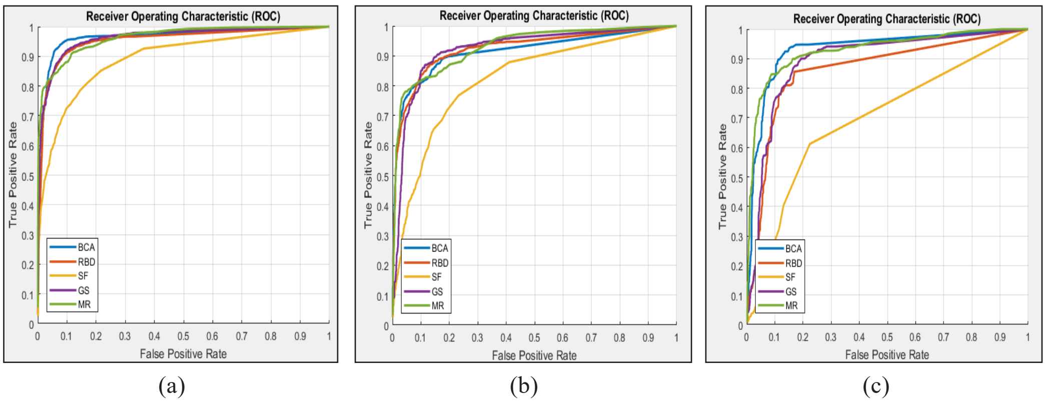
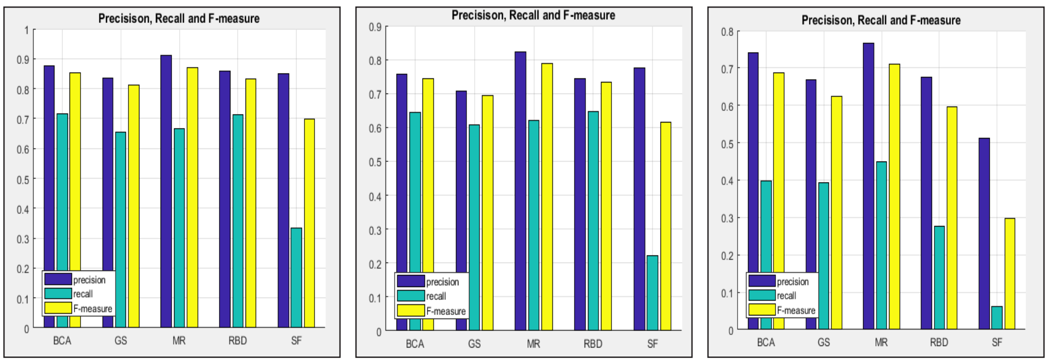


Figure 9. PR curve on the BCA comparison with other state-of-the-art models using: (a) ASD dataset; (b) ECSSD dataset; and (c) close contrast images collected from ASD dataset.

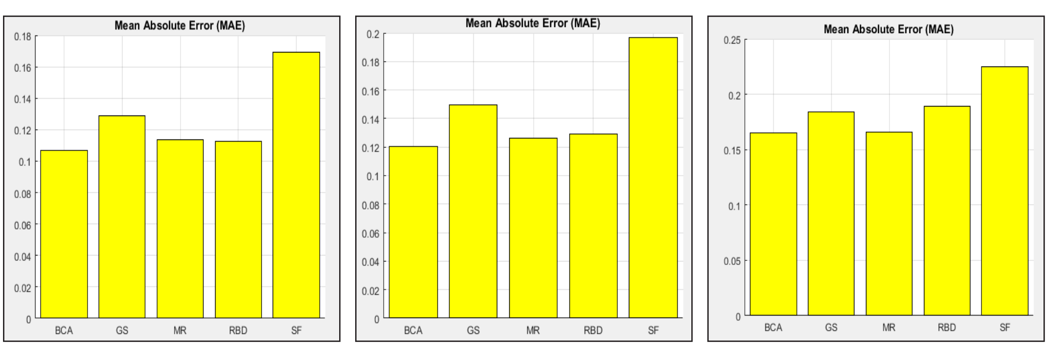
Precision, Recall and F-Measure. In the comparison made in Figure 10, the average of precision, recall and F-measure were obtained from the 3 testing images. From this histogram, the proposed model BCA had the second highest value of all after MR for all tests. Higher F-measure indicates that saliency map obtained has high value for both precision and recall.



(a) (b) (c)

Figure 10. Precision, Recall and F Measure comparison with the FBR as proposed model, using (a) ASD dataset,(b) ECSSD dataset and (c) close contrast images collected from ASD dataset.

Mean Absolute Error (MAE). Table 1 is the value of computed MAE for different test images. When compared with the RBD as a similar base algorithm assumption, BCA has been able to reduce the MAE value in the range of 3.0% to 9.4%. When comparing the MAE of BCA against other models as in Figure 11, the BCA gives the lowest MAE for all 3 test images. This indicates that the saliency map obtained is close to the ground truth. The high value of MAE shows that there are many patches on the saliency map when compared to the ground truth. Thus, it can be said that the BCA is able to produce clean background on its detection.



(a) (b) (c)

Figure 11. MAE value comparison with the proposed model, FBR using the (a) MSRA10K dataset, (b) ECSSD dataset and (c) close contrast images collected from ASD dataset.

Table 1
MAE comparison for the three testing datasets

No	Test Images Set	BCA	GS	MR	RBD	SF
1	ASD	0.10486	0.12905	0.11361	0.10807	0.16910
2	ECSSD	0.11351	0.14979	0.12648	0.12124	0.19656
3	Close Contrast Images	0.18131	0.18421	0.16610	0.20013	0.22519

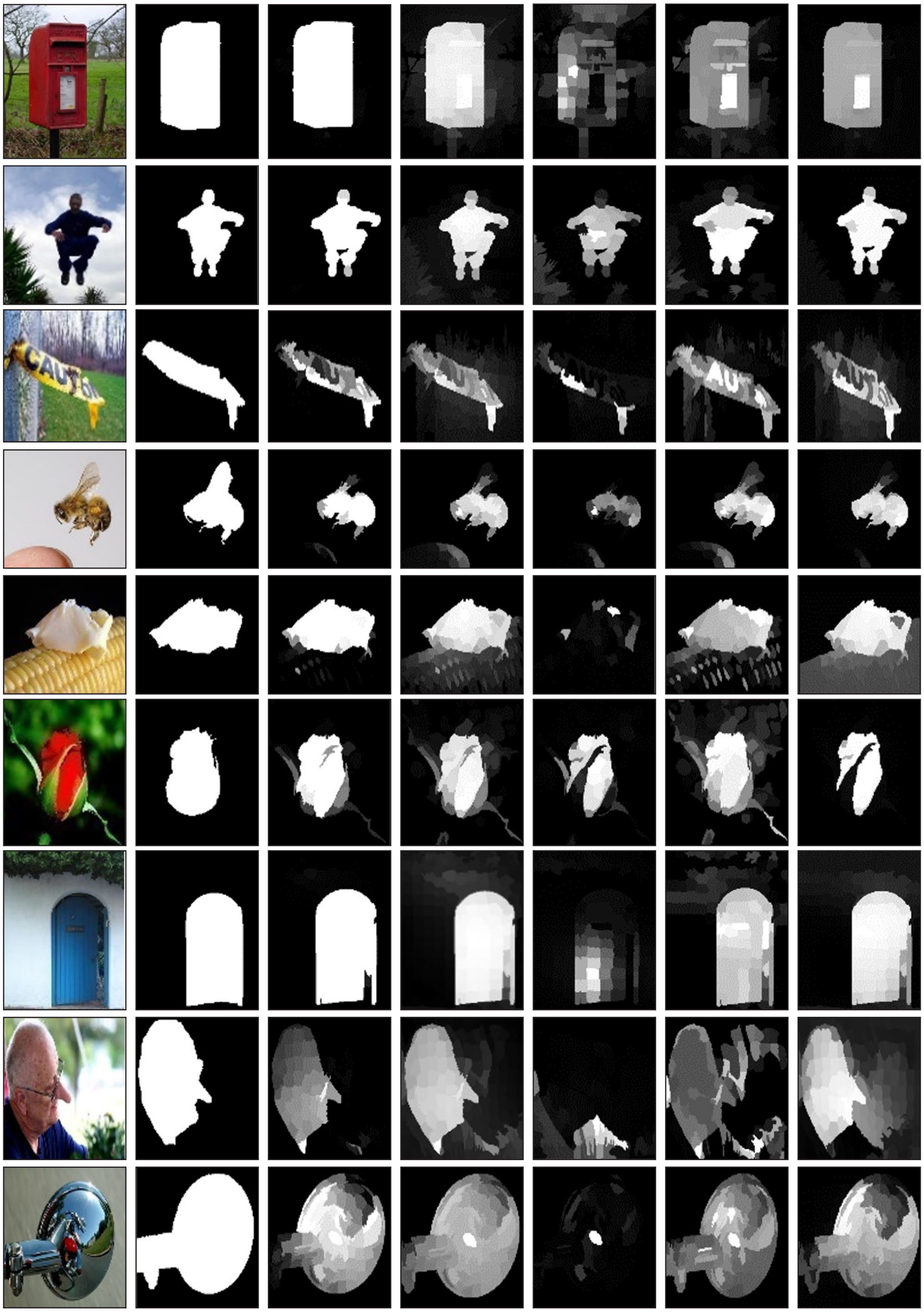
Execution Time of Salient Object Detection Comparison. The intension of salient detection is to reduce the processing time of some semantic processing purpose as the output of the detection is able to mask out the image background and only left with the most prominent object in the image. Therefore, the evaluation of detection execution time is very important and needs to be put as the result comparison. The codes of the other 4 models were accessible online as in MATLAB file (m file). They were run and compared using the MATLAB R2017b version software on a machine with the Intel i7-7500U CPU 2.70GHz, 2901 as the processor and 4G of RAM. The execution time comparison can be referred to Table 2.

Table 2
Execution time (s) for 3 different test images

No	Test Images Set	BCA	GS	MR	RBD	SF
1	ASD	0.1555	0.1315	0.1360	0.1503	0.1401
2	ECSSD	0.1466	0.1220	0.1295	0.1435	0.1299
3	Close Contrast Images	0.1579	0.1437	0.1444	0.1568	0.1535

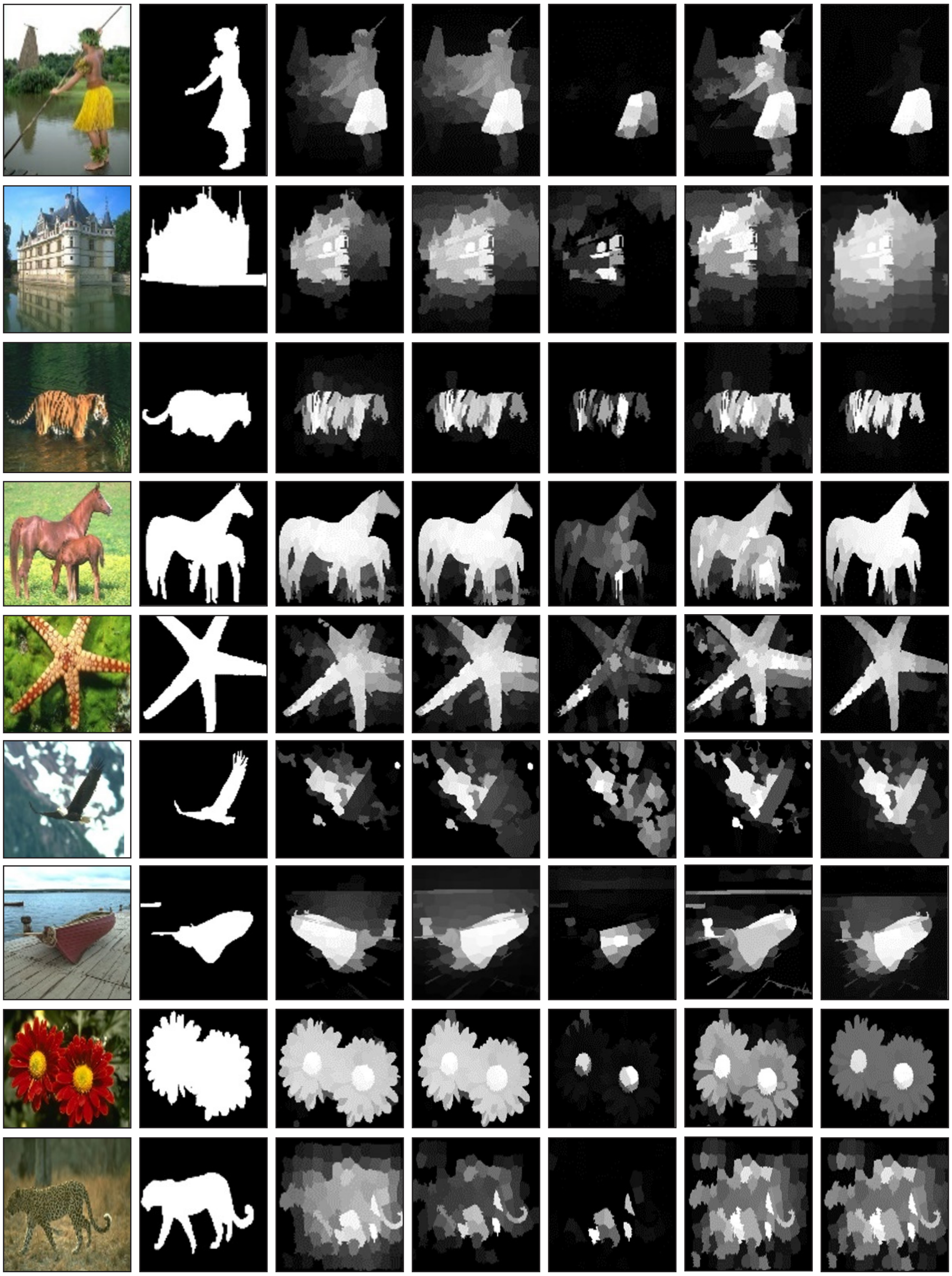
The execution time of BCA and RBD were quite consistent on all type of images. However, BCA had a slightly longer execution time in the range of 3.45% to 0.7% higher compared to RBD. This is due to additional computation that were considered in the model.

Input Images, Ground Truth and Salient Detection Comparison. The visual comparison of the saliency map obtained from BCA with other models, the BCA was able to produce clean background, resulting in minimum white patches as shown in Figures 12, 13 and 14. The results are consistent on the 3 types of test images. These are the valid reasons the BCA is able to produce the lowest MAE in previous comparison. Even though the saliency maps obtained for close contrast images do not really match the ground truth, they still produce better detections compared to the other models.



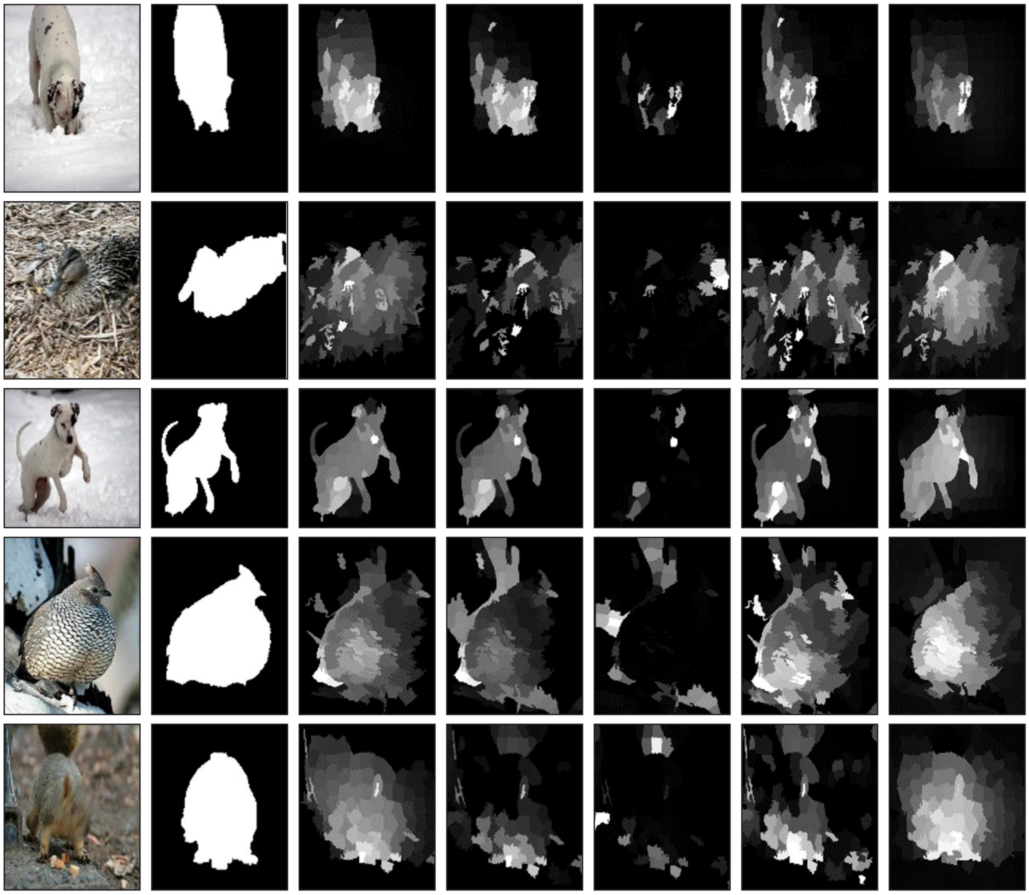
(a) (b) (c) (d) (e) (f) (g)

Figure 12. Comparison results for salient object detection images from the ASD dataset. (a) Original image, (b) Ground truth, (c) BCA (proposed), (d) RBD, (e) SF, (f) GS, (g) MR



(a) (b) (c) (d) (e) (f) (g)

Figure 13. Comparison results for salient object detection images from the ASD dataset. (a) Original image, (b) Ground truth, (c) BCA (proposed), (d) RBD, (e) SF, (f) GS, (g) MR



(a) (b) (c) (d) (e) (f) (g)

Figure 14. Comparison results for salient object detection images on close contrast image. (a) Original image, (b) Ground truth, (c) BCA (proposed), (d) RBD, (e) SF, (f) GS, (g) MR

CONCLUSION

This paper has presented an improved salient object detection model using combination features known as BCA on its saliency computation. The model has overcome the problem where high contrast background patches are falsely assigned as the foreground. On the other hand, the detection on close contrast image has been improved with lesser false positive rate. With the lowest MAE obtained, it indicates clear background with less white patches successfully obtained. As the lowest MAE is obtained for the close contrast image, it indicates that the proposed model is suitable to be applied for the salient object detection on that kind of image. The integrated features consisting of contrast, spatial location, force interaction and boundary ratio have been able to produce a novel equation of boundary connectivity. This new consideration has resulted in improved saliency map and has outperformed the referred model, RBD in most of the performance comparisons. Despite that, the model has embedded the force interaction feature that has not previously been applied in any existing salient detection models. Therefore, the combination features in BCA has been successful in producing better salient object detection by its background measure computation. Despite of the improvement on the result of detections, the BCA execution time requires a longer duration due to the additional feature extraction process. This weakness could become the component that can be improved in the future. Furthermore, the learning technique could be integrated with the model to improve detection especially for close contrast images.

ACKNOWLEDGEMENT

We would like to thank to the anonymous reviewers for their constructive comments and opinions that helped us in producing significant improvement on this research paper.

REFERENCES

Achanta, R., Estrada, F., Wils, P., & Süsstrunk, S. (2008). Salient region detection and segmentation. In A. Gasteratos, M. Vincze, & J. K. Tsotsos (Eds.), *Computer vision systems* (pp. 66-75). Heidelberg, Germany: Springer.

Achanta, R., Hemami, S., Estrada, F., Sabine, S., & Epfl, D. L. (2009, June 20-25). Frequency-tuned salient region detection. In *2009 IEEE Conference on Computer Vision and Pattern Recognition* (pp. 1597-1604). Miami Beach, Florida.

Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., & Susstrunk, S. (2012). SLIC Superpixels Compared to State-of-the-Art Superpixel Methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 34*(11), 2274-2282.

Ahn, E., Kim, J., Bi, L., Kumar, A., Li, C., Fulham, M., & Feng, D. D. (2017). Saliency-based lesion segmentation via background detection in dermoscopic images. *IEEE Journal of Biomedical and Health Informatics*, *21*(6), 1685-1693.

Baluja, S. (1997). Expectation-based selective attention for visual monitoring and control of a robot vehicle. *Robotics and Autonomous Systems*, *22*(3-4), 329-344.

Borji, A., Cheng, M. M., Jiang, H., & Li, J. (2014). Salient object detection: a survey. *Computattional Visual Media, 5*(2), 117-150.

Borji, A., Cheng, M. M., Jiang, H., & Li, J. (2015). salient object detection: a benchmark. *Image Processing, IEEE Transactions On*, *24*(12), 5706-5722.

Bucha, V., Uchida, S., & Ablameyko, S. (2006). Interactive road extraction with pixel force fields. *Proceedings-International Conference on Pattern Recognition*, *4*(i), 829-832.

Bucha, V., Uchida, S., & Ablameyko, S. (2007). Image pixel force fields and their application for color map vectorisation. *Proceedings of the International Conference on Document Analysis and Recognition, ICDAR*, *2*, 1228-1232.

Cheng, M., Zhang, G., Mitra, N. J., Huang, X., & Hu, S. (2011). cn-global contrast based salient region detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence 2011*, *37*(3), 409-416.

Chou, Y., & Su, P. C. (2016, May 27-29). Efficient content-based cropping using visual saliency and SLIC for image retargeting. In *2016 IEEE International Conference on Consumer Electronics* (pp. 1-2). Nantou, Taiwan.

Ghayoor, A., Sadri, A., & Shirazi, A. A. B. (2010, May 26-28). Image registration method based on physical forces for color images. In *2010 Fourth Asia International Conference on Mathematical/Analytical Modelling and Computer Simulation* (pp. 350-353). Bornea, Malaysia.

Guo, C., & Zhang, L. (2010). A novel multiresolution spatiotemporal saliency detection model and its applications in image and video compression. *IEEE Transactions on Image Processing*, *19*(1), 185-198.

Hu, P., Wang, W., Zhang, C., & Lu, K. (2016). Detecting salient objects via color and texture compactness hypotheses. *IEEE Transactions on Image Processing*, *25*(10), 4653-4664.

Hurley, D. J., Nixon, M. S., & Carter, J. N. (1999). Force field energy functionals for image feature extraction. *Image and Vision Computing*, *20*(5-6), 311-317.

Itti, L. (2004). Automatic foveation for video compression using a neurobiological modelof visual attention. *IEEE Transactions on Image Processing, 13*(10), 1304-1318.

Itti, L., Koch, C., & Niebur, E. (2001). A Model of Saliency-Based Viusla Attention for Rapid Scene Analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *23*(6), 674-680.

Jiang, H., Wang, J., Yuan, Z., Liu, T., Zheng, N., & Li, S. (2011, August 29-September 2). Automatic salient object segmentation based on context and shape prior. In *Proceeding of British Machine Vision Conference* (pp. 1-12). University of Dundee, Scotland.

Jiang, L., Koch, A., & Zell, A. (2015). Salient regions detection for indoor robots using RGB-D data. *Proceedings-IEEE International Conference on Robotics and Automation*, *2015–June*(June), 1323-1328.

Jing, H., He, X., Han, Q., & Niu, X. (2014). Neurocomputing. *Neurocomputing*, *124*, 57-62.

Katsuki, F., & Constantinidis, C. (2014). Bottom-Up and Top-Down Attention : Different Processes and Overlapping Neural Systems. *The Neuroscientist, 20*(5), 509-521.

Koch, C., & Ullman, S. (1985). Shifts in selective visual attention: towards the underlying neural circuitry. *Human Neurobiology, 4*(4), 219-227.

Li, C., Huo, H., & Fang, T. (2016, July 11-12). Oil depots detection from high resolution remote sensing images. In *2016 International Conference on Audio, Language and Image Processing (ICALIP)* (pp. 285-288). Shanghai, China.

Li, J., Member, S., & Levine, M. D. (2013). Visual saliency based on scale-space analysis in the frequency domain. *IEEE transactions on pattern analysis and machine intelligence, 35*(4), 996-1010.

Li, S., Xu, M., Ren, Y., & Wang, Z. (2017). Closed-form optimization on saliency-guided image compression for HEVC-MSP. *IEEE Transactions on Multimedia*, *20*(1), 155-170.

Liu, D., Cong, M., Du, Y., & Gao, S. (2013, December 12-14). Robot behavior selection using salient landmarks and object-based attention. In *2013 IEEE International Conference on Robotics and Biomimetics, ROBIO 2013* (pp. 1101-1106). Shenzhen, China.

Liu, Y., Zhang, J., Tjondronegoro, D., Geva, S., & Li, Z. (2008, November 26-28). An improved image segmentation algorithm for salient object detection. In *2008 23rd International Conference Image and Vision Computing New Zealand* (pp. 1-6). Christchurch, New Zealand.

Luo, Y., Wang, P., Li, W., Shang, X., & Qiao, H. (2016, June 12-15). Salient object detection based on boundary contrast with regularized manifold ranking. In *Proceedings of the World Congress on Intelligent Control and Automation (WCICA)* (pp. 2074-2079). Guilin, China.

Luz, G., Ascenso, J., Brites, C., & Pereira, F. (2017, October 16-18). Saliency-driven omnidirectional imaging adaptive coding: Modeling and assessment. In *2017 IEEE 19th International Workshop On Multimedia Signal Processing (MMSP)* (pp. 1-6). Luton, UK.

Manke, R., & Jalal, A. S. (2016, December 3-4). Salient region detection using fusion of image contrast and boundary information. In *2016 11th International Conference on Industrial and Information Systems (ICIIS)* (pp. 199-202). Roorkee, India.

Milanese, R., Gil, S., & Pun, T. (1995). Attentive mechanisms for dynamic and static scene analysis. *Optical Engineering*, *34*(8), 2428-2434.

Niu, Y. (2018). Salient object segmentation based on superpixel and background connectivity prior. *IEEE Access*, *6*, 56170-56183.

Pinto, Y., van der Leij, A. R., Sligte, I. G., Lamme, V. A., & Scholte, H. S. (2013). Bottom-up and top-down attention are independent. *Journal of vision, 13*(3), 16-16.

Perazzi, F., & Kr, P. (2012, June 16-21). Saliency filters Contrast based filtering for salient region detection. In *2012 IEEE Conference on Computer Vision and Pattern Recognition* (pp. 733-740). Providence, RI, USA.

Pritch, Y., Sorkine-hornung, A., Mangold, S., & Gross, T. (2013, September 15-18). Content-aware compression using saliency-driven image retargeting. In *2013 IEEE International Conference on Image Processing* (pp. 1845-1849). Melbourne, VIC, Australia.

Srivastava, S., Mukherjee, P., & Lall, B. (2016, June 12-15). Adaptive Image Compression Using Saliency and KAZE Features. In *2016 International Conference on Signal Processing and Communications (SPCOM)* (pp. 1-5). Bangalore, India.

Tang, L., Li, H., Wu, Q., & Ngan, K. N. (2018). Boundary guided optimization framework for saliency refinement. *IEEE Signal Processing Letters*, *9908*(c), 1-1.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*(1), 97-136.

Wang, J., Zhang, K., Madani, K., Sabourin, C., & Zhang, J. (2015, July 21-23). Salient foreground object detection based on sparse reconstruction for artificial awareness. In *2015 12th International Conference on Informatics in Control, Automation and Robotics (ICINCO)* (Vol. 2, pp. 430-437). Colmar, France.

Wei, Y., Wen, F., Zhu, W., & Sun, J. (2012). Geodesic saliency using background priors. In *European conference on computer vision* (pp. 29-42). Heidelberg, Germany: Springer.

Woodbridge, J., Lan, M., Sarrafzadeh, M., & Bui, A. (2011, July 26-29). Salient segmentation of medical time series signals. In *2011 First IEEE International Conference on Healthcare Informatics, Imaging and Systems Biology (HISB 2011)* (pp. 1-8). San Jose, CA, USA.

Wu, P. H., Chen, C. C., Ding, J. J., Hsu, C. Y., & Huang, Y. W. (2013). Salient region detection improved by principle component analysis and boundary information. *Ieee Transactions on Image Processing*, *22*(9), 3614-3624.

Xiao, R., & Yeh, M. (2017, June 12-14). A new method with saliency detection for image quality assessment. In *2017 IEEE International Conference on Consumer Electronics* (pp. 75-76). Taipei, Taiwan.

Xiuli, T., Shuhan, C., & Xuelong, H. (2017, October 20-22). Salient object detection by combing eye fixation prediction and semantic segmentation. In *2017 13th IEEE International Conference on Electronic Measurement Instruments (ICEMI)* (pp. 407-411). Yangzhou, China.

Yan, Q., Xu, L., Shi, J., & Jia, J. (2013, June 23-28). Hierarchical saliency detection. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (pp. 1155-1162). Portland, OR, USA.

Yang, B., Zhang, X., Chen, L., Yang, H., & Gao, Z. (2017). Edge guided salient object detection. *Neurocomputing*, *221*(March 2016), 60-71.

Yang, C., Zhang, L., Lu, H., Ruan, X., & Yang, M. H. (2013, June 23-28). Saliency detection via graph-based manifold ranking. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (pp. 3166-3173). Portland, Oregon.

Zhang, H., Goldman, S. A., & Louis, S. (2006, October 8-11). Image Segmentation Using Salient Points-Based Object Templates. In *2006 International Conference on Image Processing, Atlanta* (pp. 765-768). Atlanta, GA, USA.

Zhang, L., Lv, X., Chen, J., & Zhang, L. (2016, July 10-15). Region of interest detection based on salient feature clustering for remote sensing images. In *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)* (pp. 88-91). Beijing, China.

Zhang, Q., Lin, J., & Li, X. (2016). Salient object detection via structure extraction and region contrast. *Journal of Information Science and Engineering*, *32*(6), 1435-1454.

Zhang, Q., Lin, J., Tao, Y., Li, W., & Shi, Y. (2017). Neurocomputing salient object detection via color and texture cues. *Neurocomputing*, *243*, 35-48.

Zhou, L., Chen, Y., & Yang, Z. (2017). Foreground and background propagation based salient region detection. *2016 9th International Symposium on Computational Intelligence and Design (ISCID)*, *1*, 110-113.

Zhu, W., Liang, S., Wei, Y., & Sun, J. (2014, June 23-28). Saliency optimization from robust background detection. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (pp. 2814-2821). Columbus, Ohio.

Pertanika J. Sci. & Technol. 27 (4): 1759 - 1781 (2019)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Detection and Recognition via Adaptive Binarization and Fuzzy Clustering

Saad Mohmad Saad Ismail, Siti Norul Huda Sheikh Abdullah and Fariza Fauzi*

Center of Cyber Security, Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, 43600 UKM, Bangi, Selangor, Malaysia

ABSTRACT

Detection and identification of text in natural scene images pose major challenges: image quality varies as scenes are taken under different conditions (lighting, angle and resolution) and the contained text entities can be in any form (size, style and orientation). In this paper, a robust approach is proposed to localize, extract and recognize scene texts of different sizes, fonts and orientations from images of varying quality. The proposed method consists of the following steps: preprocessing and enhancement of input image using the National Television System Committee (NTSC) color mapping and the contrast enhancement via mean histogram stretching; candidate text regions detection using hybrid adaptive segmentation and fuzzy c-means clustering techniques; a two-stage text extraction from the candidate text regions to filter out false text regions include local character filtering according to a rule-based approach using shape and statistical features and text region filtering via stroke width transform (SWT); and finally, text recognition using Tesseract OCR engine. The proposed method was evaluated using two benchmark datasets: ICDAR2013 and KAIST image datasets. The proposed method effectively dealt with complex scene images containing texts of various font sizes, colors, and orientation; and outperformed state-of-the-art methods, achieving >80% in both precision and recall measures.

Keywords: Adaptive binarization; fuzzy C-means; image enhancement; statistical and geometrical features; text detection; text extraction

ARTICLE INFO

Article history:
Received: 30 January 2019
Accepted: 03 July 2019
Published: 21 October 2019

E-mail addresses:
ababnah2010@gmail.com (Saad Mohmad Saad Ismail)
snhsabdullah@ukm.edu.my (Siti Norul Huda Sheikh Abdullah)
fariza.fauzi@ukm.edu.my (Fariza Fauzi)
* Corresponding author

INTRODUCTION

With the growing popularity of mobile/ wearable devices with embedded digital cameras, an abundance of scene images is available. It is common to find text present/

ISSN: 0128-7680
e-ISSN: 2231-8526

embedded in scene images (scene text) such as shop signage, street names and billboards which can contribute significant information about a particular scene. Text has been and still is the most intuitive way for most people to represent information. As such, extracting text from images of real scenes has become an important task especially with the advent of life changing technologies such as the Internet of Things, robotics, automation and augmented reality. Scene text detection and extraction play an important role in many applications such as detecting texts on signposts (Raghunandan et al., 2018), billboards, nameplates, labels, license plate recognition (Huang et al., 2018; Li et al., 2018), invoice number detections (Sun et al., 2019) , mobile text reader, ticket number reader (Kraisin & Kaothanthong, 2018), handphone number reader, scene-to-text conversion and pronunciation for visually impaired people (Qaisar et al., 2019).

Natural scene images are usually taken using digital and mobile cameras. These photos are diverse, of varying quality (Nagaraju et al., 2015; Risnumawan et al., 2014) making the detection and extraction of scene text a difficult and challenging task. Various issues include changes in light intensity, text alignment, font size, color, camera angles, occlusion and so on (Behzadi & Safabakhsh, 2018; Ning et al., 2015; Raghunandan et al., 2018; Wang et al., 2015; Zhang et al., 2013). Generally, text extraction and recognition techniques can automatically detect areas of texts in a scene image by marking boundaries (usually bounding boxes) and reading its text content as a series of (Unicode) characters. The digitized text can then be further processed by a computer. Extracting text from real scene images generally consists of three phases: (1) detection and localization, (2) enhancement and segmentation, and (3) optical character recognition (OCR). The most difficult problem when dealing with natural text extraction and recognition is due to the differences in the font, color, alignment of text size, change in lighting, and reflections (Jadhav et al.,2013; Sulaiman et al., 2019). These persistent challenges have made histogram-based segmentation methods (Behzadi & Safabakhsh, 2018; Ning et al., 2015; Raghunandan et al., 2018; Wang et al., 2015; Zhang et al., 2013) and handcrafted approaches to be trivial and insignificant in such scenario or situation.

Many researchers have proposed several text extraction methods from complex images (Sumathi et al., 2012). Although each method has its own unique approach and objectives, they have one common goal which is to obtain the best extraction of text from an image in order to maximize text recognition. An in-depth literature review on OCR has been presented in (Jung et al., 2004; Zhang et al., 2013) whereby each review highlighted its own research gaps, benchmark dataset, evaluation criteria, drawbacks and future direction. Jung et al. (2004) gave a comprehensive survey in classifying algorithms proposed to address the related problems in text extraction. A typical text extraction process involves detection, localization, tracking, extraction, enhancement, and recognition of the text from any scene image. Karanje and Dagade (2014) focused on the advantages and disadvantages of the major categories of text extraction approaches. They also described state-of-the-art methods for the detection of text, text segmentation, and character recognition in natural scene images. Samadhiya and Khatri (2014) reviewed various text extraction algorithms and discussed the performance evaluation and challenges. The authors offered a collection of recent techniques for text information extraction.

Besides that, several researches implemented the K-means clustering algorithm to perform segmentation and recognition. Burney and Tariq (2014) proposed K-means clustering for segmentation where dissimilar pixels from the source image were divided into several (or a predetermined number) similar regions or segments. K-means clustering can also be used directly for classification. It distinguishes the objects regions into text or non-text (Wang et al., 2011). On the other hand, the drawback of K-means clustering approach in text extraction and segmentation is that it does not perform well when dealing with natural image pixels that have very low or very high contrast. This weakness causes the separation of text from its background to become nontrivial.

Shivakumara et al. (2011) described a Laplacian operator based on the method of frequency domain model. In this approach, the input image was filtered using Fourier-Laplacian transform. Then it used K-means clustering to identify regions of candidate text on the basis of the maximum difference. To overcome the limitation of K-means, a new iterative nearest neighbor symmetry has been introduced intentionally for restoring missing text whereas fixing window based on angular relationship relying on sub-bands and its fused bands is proposed to improve arbitrary oriented text in the natural scene environment (Raghunandan et al. 2018). Apart from classic natural scene-handcrafted methods, several researchers were also discovering machine-crafted natural text segmentation by introducing word-fence and fully convolutional Densenet (Behzadi & Safabakhsh, 2018) whilst Huang et al. (2018) presented a unified end-to-end trainable deep network, which could simultaneously locate and recognize vessel plate number.

Phan et al. (2009) developed an approach to detect text with the Laplacian operator. Then K-means is used to classify all pixels in clusters. Based on the literature, less attention was given to other clustering approaches, such as fuzzy c-means (FCM) clustering. FCM is a frequently used clustering method in pattern recognition which allows one piece of data to belong to two or more clusters (Horvath, 2006; Rajaby et al., 2016). As such, it is investigated in this paper for the tasks of text detection and extraction.

The remainder of the paper is organized as follows: The Method section describes the overall proposed method consisting of the preprocessing and enhancement phase, followed by the hybrid adaptive segmentation and clustering phase to detect, extract and recognize text. The Results and Discussion section examines measurements of extraction and recognition of different datasets that are used to objectively compare the proposed method to existing text extraction methods. Experimentally, these different methods are compared using different datasets with different environment types. Finally, a conclusion is drawn.

METHOD

The block diagram of the proposed method is illustrated in Figure 1. Given an input image, text color image remapping was performed using NTSC color distance and histogram normalization to obtain a clean grayscale image with clear and high contrasting background and text, addressing the poor/low contrast issue for the effective recognition. Then, to convert the clean grayscale image into a binary image, a hybrid approach of clustering and adaptive binarization was performed. The clustering of grayscale image into two clusters was achieved using FCM. The output was combined with the outputs of the adaptive binarization of color image channels and the complement of the binary image. The aim of this step is to ensure the conversion is performed without fading any text regions with high appearance, in addition to unifying the background for all text regions within the same image. Finally, the regions were extracted using a connected component technique based on several statistical distributions and shape features. The extracted text regions were filtered using a rule-based technique and stroke width transform using Manhattan distance transform. For character recognition, the open-source Tesseract OCR engine was employed. The subsequent sections will detail out each stage in the proposed method.

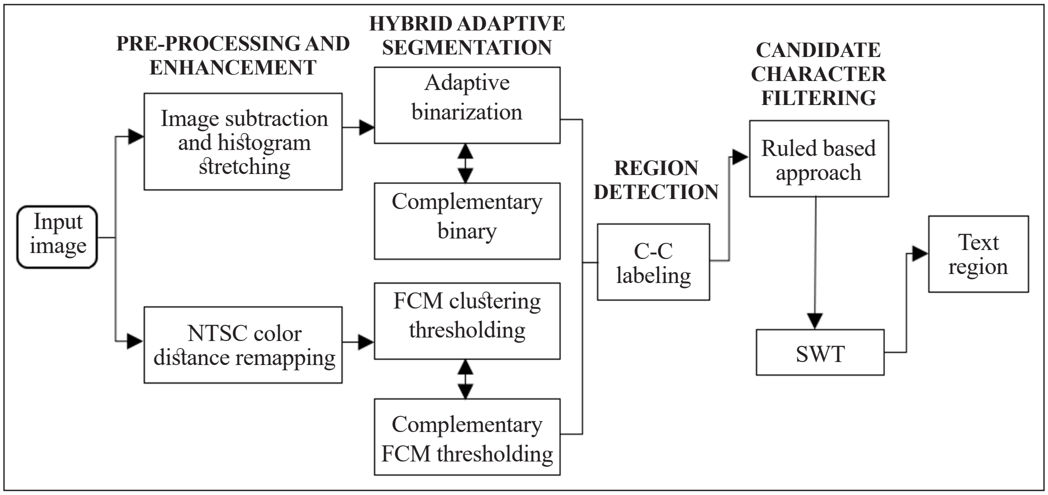


Figure 1. Block diagram of the proposed approach

Preprocessing and Enhancement

Natural scene images can contain noise for example, salt and pepper noise, impulse noise etc. or they can be blurred due to inadvertent movement of camera, poor lighting condition causing over- or under-exposure, light reflection and non-uniform color. For that purpose, image preprocessing and enhancement techniques are applied to reduce such noise. In the proposed framework, color distance remapping and low contrast enhancement by subtraction and normalization using histogram stretching are performed.

Color Space Mapping. The proposed image color enhancement method is aimed to suppress the reflective light, uneven illumination in noisy scene images. Hence, the source color image is enhanced by using the proposed exploited NTSC color mapping method. In general, there are different image color spaces such as RGB, NTSC, HSV and each one has its own unique features. Here, the NTSC color space was selected primarily for its ability to characterize (Shah & Thakar, 2010) and separate gray color from other colors (red and green). Ultimately, this exploited remapping process can easily tackle the aforementioned problems.

In the NTSC format, image data consists of three components: luminance (Y), and chrominance: hue (I), and saturation (Q). The first component, luminance, represents the grayscale information, while the last two components represent chrominance (color information). Usually, background pixel values are mostly classified in luminance component. By multiplying these values with a small value, $\propto$, the gray (luminance) values decreases to merely zero or black. Contrariwise, multiplying the chrominance by big values (β,γ) the appearance of foreground (object) increases to merely white. Heuristically, the best $\propto$,β,γ values are obtained. Equation 1 denotes the NTSC color distance remapping function where luminance and chrominance are multiplied by $\propto$,β,γ.

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.596 & 0.211 \\ 0.587 & -0.274 & -0.523 \\ 0.114 & -0.322 & 0.312 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \times \begin{bmatrix} \propto \\ \beta \\ \gamma \end{bmatrix}, \quad (1)$$

where: $\propto$ =0.6, β=8, γ=16.

This mapping process solves the light reflection and low contrast flaws in scene image. One main benefit of this format is that the grayscale information is separated from color data, therefore, both color as well as black and white sets can share the same channel.

Figure 2 illustrates the result of the proposed NTSC color mapping method: the original image, shown in Figure 2(a) is compared to images in Figure 2(b)-2(d) which are processed using regular grayscale conversation / RGB channels and images in Figure 2(e)-2(g) are processed using the proposed NTSC color distance remapping function. As demonstrated, the proposed method effectively differentiated the text from background, resulting in a cleaner image obtained from the scene image in Figure 2(e)-2(g), with the exceptions of scene images having a very low contrast with a high light reflection at the same time, or when there are highly-effective shadows in the text region. Ultimately, the proposed NTSC enhancement step addresses the low contrast and illumination problems, providing cleaner images to the subsequent step (Hybrid Adaptive Segmentation) of the overall method which aims to efficiently detect and extract text from scene images.

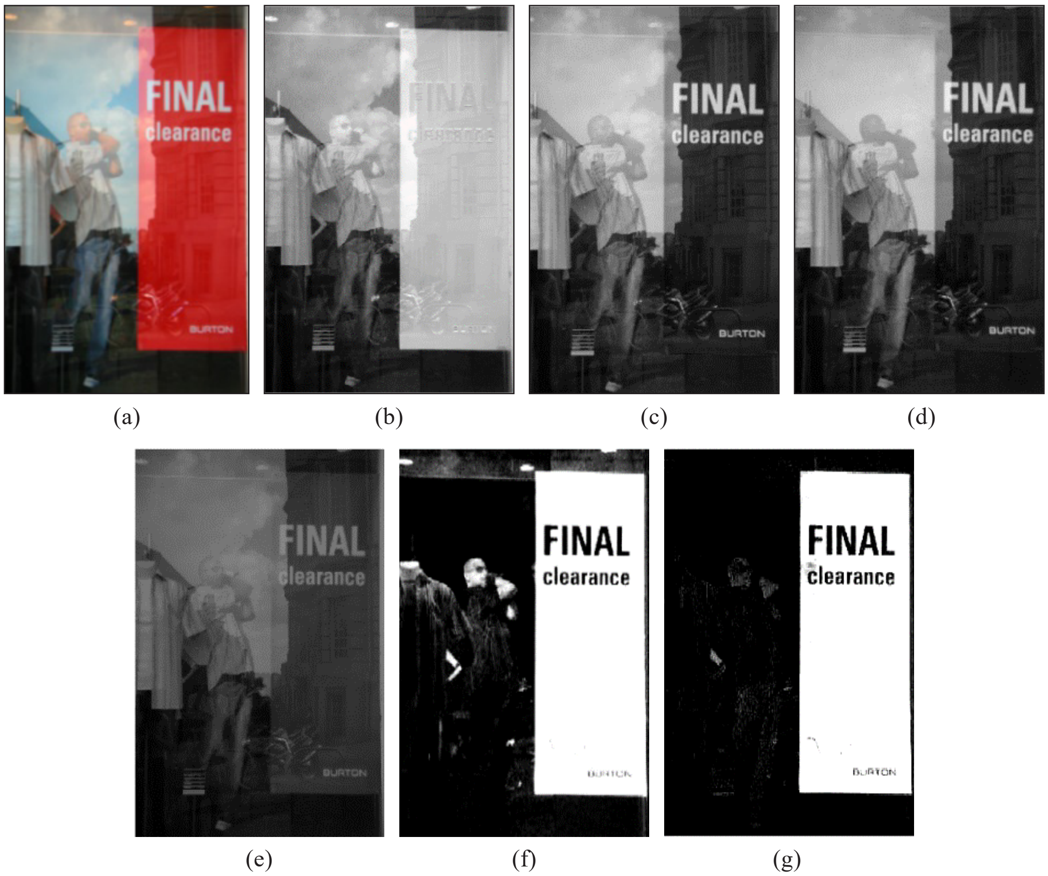


Figure 2. An example showing the effectiveness of the proposed image enhancement. method: (a) original image (b-d) before and (e-g) after introduction of the proposed exploited NTSC for image enhancement method according to RGB channels (red, green, blue, respectively).

Hybrid Adaptive Segmentation and Clustering

The proposed hybrid adaptive segmentation method combines the results of two segmentation techniques: FCM clustering-based segmentation and adaptive binarization-based segmentation. The hybrid approach draws on the strengths of FCM clustering in segmentation and adaptive binarization in addressing low contrast and illumination problems, which is essentially a drawback in FCM clustering. This results in an improved segmentation method which retains most of the text regions with minimal noise.

Segmentation using Proposed Fuzzy C-Means (FCM) Clustering. For text extraction from scene images, the FCM clustering was utilized for binarization purposes. First, the grayscale image was clustered into two categories, and then a global threshold method was applied to set pixels in each cluster to either black or white. The binarization process utilizing FCM clustering is described in Algorithm 1 which was used in Horvath (2006).

The algorithm assumes the scene image has been converted into the grayscale feature space with C number of clusters and the stop condition and fuzziness parameter are denoted by ε and M respectively. Here, it is used for image binarization, hence C was set to 2 to segment out the text from the background.

Algorithm 1. The Proposed FCM Clustering Method for Image Binarization

Input: Grey Image
Output: Binary Image
Start
Step 1: Cluster image in the grayscale feature space, with the following conditions: number of clusters = C, fuzziness index = f, stop condition = ε.
Step 2: Repeat for each pixel $I(i,j)$ of image I.
Step 3: Compute which cluster C does pixel $I(i,j)$ belong to.
Step 4: Identify if there exist a segment R_k whose points belong to the same cluster C, in the closest surroundings of pixel $I(i,j)$.
Step 5: If segment R_k exists, then pixel $I(i,j)$ is added to segment R_k, else create a new segment R_n and add pixel $I(i,j)$ to the new segment R_n.
Step 6: Merge all segments which belong to one cluster and are neighbors.
Step 7: Arrange borders of all segments.
Step 8: The pixels belonging to cluster one are set to white and the pixels belonging to cluster two are set to black
End

The experiment results indicated that the FCM clustering technique gave good results for many cases. However, the drawback of this method is that it is affected by strong lighting and non-uniform text color that cannot be overcome. Figure 3 presents an example of the binary segmentation results using FCM clustering algorithm.

Segmentation using Proposed Adaptive Binarization. Adaptive compound binarization is the second segmentation technique proposed to segment and binarize text from scene images. The adaptive binarization is expressed in the equation below:

$$T_{Adaptive} = \mu_l * (1 - T_{G\ Dynamic}) \qquad (2)$$

where $T_{Adaptive}$ is the local threshold value and μ_l is the local mean value of integral image pixels estimated based on a dynamic generation window size ($T_{G\ Dynamic}$). Based on this $T_{Adaptive}$ value, the binarization process is defined in equation (3). With the assumption that I_{binary} *(x, y)* is the binary image and *I(x, y)* is the grayscale image, therefore when the current pixel *I(x, y)* is less than $T_{Adaptive}$, it is set to black, otherwise it is set to white.

$$T_{binary}(x,y) = \begin{cases} \text{black, } I(x,y) & < T_{Adaptive} \\ \textit{white} & \textit{otherwise} \end{cases} \quad (3)$$

As mentioned in the beginning of the section, the proposed adaptive binarization compensates the drawbacks of FCM clustering method wherein its poor ability to segment images for very low contrast or high illuminance e.g. texts “A 12” and “A 120” appear as a black box, “PR” is missing, “I” is broken and missing text “inhabited environment” in Figure 3(b). Figure 3(c) presents the results of the proposed adaptive compound binarization method on the original set of images in Figure 3(a). It is demonstrated in Figure 3(c) that adaptive binarization produces better results in extracting text from low contrast or high illuminance scene images. However, it is also observed that many unnecessary background details (noise) are also present. Hence, a hybrid adaptive segmentation step is recommended in the overall proposed approach for text detection and recognition from degraded scene images.

(a) (b) (c)

Figure 3. Sample segmentation results. (a) Original image set (b) segmentation results using FCM clustering and (c) segmentation results using adaptive compound binarization.

Hybrid Segmentation. Final segmentation was the aggregated results of FCM clustering for the grayscale image and its compliment with the results of the proposed adaptive binarization of the grayscale image and its compliment. This process gives the best segmentation that is capable of preserving text region in scene images with lowest noise. The following equation denotes the hybrid segmentation process.

$$I_{hybrid} = I_{binary} + I^{c}_{binary} + U + U^{c} \tag{4}$$

where I_{binary} is the adaptive binarization image c.f. equation (3), I^{c}_{binary} is the complement adaptive binarization image, U is the FCM binary image and U^{c} the complement FCM binary image. Figure 4 shows some example results of the proposed hybrid segmentation step.

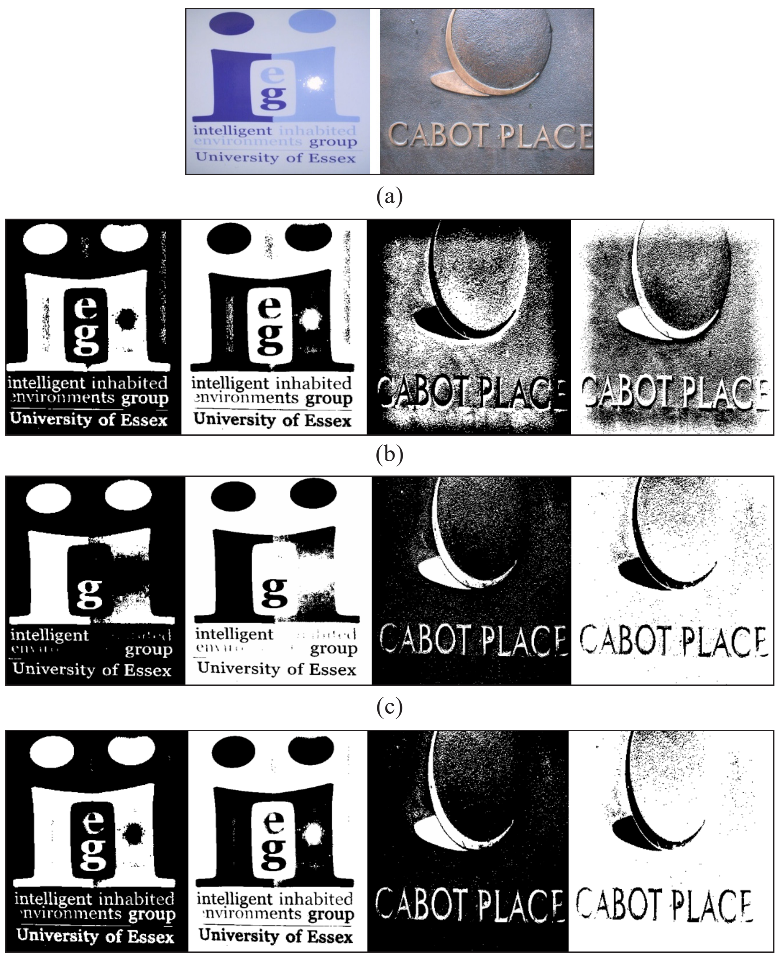


Figure 4. Example results of the proposed hybrid segmentation method: (a) original images (b) adaptive compound binarization and its complement (c) FCM segmentation and its complement (d) Final segmentation for the original images and their complement.

Text Extraction from Binary Image

Preserving all the text information together with other objects as the foreground is essential. After applying the proposed hybrid segmentation step, a binary image was produced. All the foreground objects were then, labelled using connected component method (Nosal, 2008). The challenge now was to remove the non-text objects from the foreground in the text extraction phase. The text extraction phase was divided into two stages: local character and text region filtering. In the first stage, a rule-based local character filtering was implemented using shape features whereas text region filtering used Stroke Width Transforms (SWT) to eliminate non-text objects.

CC-based methods typically use a bottom-up approach by grouping small components into successively larger components until all regions are identified in the image (Kumar et al., 2010; Raj & Ghosh, 2014; Yao et al., 2007). Statistical and geometrical analyses are required to merge the text components based on the spatial arrangement of those components. In the proposed method, seven geometric and shape features, tabulated in Table 1, are considered in the overall proposed detection and extraction method for localizing the text regions.

Table 1
List of features used in the rule-based local character filtering for text localization

Feature	Definition
Eccentricity of the region	Eccentricity = $\sqrt[2]{1-(\frac{a}{b})^2}$
Relative Smoothness	$1-\frac{1}{1+\sigma^2}$, $\boldsymbol{\sigma}$ is the standard deviation
Area	$Area\ ratio = \dfrac{Region\ pixels}{Total\ image\ pixels}$
Perimeter	Perimeter = $\sum_{i=0}^{m-1}$ Length c_i, $c_i = (0 \ldots m-1)$
Length and Width of the Region	Area high and width size
Aspect ratio	The ratio of pixels in the region calculated by dividing the width by the height of the region.
Orientation	Angle Between the x Axis and Major Axis of the Region

Proposed rule-based Approach for Local Character Filtering. For text detection and extraction, a set of hierarchical rules was developed, in relation to the above-mentioned geometric features to filter out the non-text areas from the CC regions or CC listing as described in Algorithm 1 The degree of variability for each feature is shown in Table 2 and Table 3.

Table 2
Degree of geometrical and shape features used for constructing the local character filtering rules

No.	**Feature**	**Low**	**Normal**	**High**
1	Eccentricity	<0.2	0.2 to 0.6	>0.8
2	Area	<70	70 to 2000	>10000
3	Relative Smoothness	<0.01	0.01 to 0.06	>0.08
4	Perimeter	<20	20 to 100	>200
5	Length and Width	<50	50 to 200	>200
6	Aspect ratio	<1	1 to 5	>7
7	Orientation	0^0 to 35^0	-90^0 to 35^0	70^0 to 90^0

*Please take note that for each row vertical represents conjunction (AND operator) and horizontal represents conjunction (OR operator) and between rows represents (OR operator)

Table 3
Decision table for non-character region

No.	Feature	Condition		Decision
		Low	High	
1	Eccentricity		1	Rule 1
2	Area	1	1	
3	Relative Smoothness		2	Rule 2
4	Orientation	2		
5	Length and Width	3	3	Rule 3
6	Aspect ratio		4	Rule 4
7	Perimeter	4	4	

*Please take note that for each row vertical represents conjunction (AND operator) and horizontal represents conjunction (OR operator) and between rows represents (OR operator)

Algorithm 2: The proposed rule-based approach for text regions filtering

Input: Child binary image of labelled component
Output: character or non-character regions
Begin
 Step 1: if (Eccentricity = high **and** Area = low **or** high) **or** *//Rule 1*
 Step 2: (Relative Smoothness = high **and** the Orientation = low) **or** *//Rule 2*
 Step 3: **(**Length and Width = high **or** low) **or** *//Rule 3*
 Step 4: **(**Aspect ratio = high **and** Perimeter = high **or** low) *//Rule 4*
 Step 5: **then** Region = non-character
 Step 6: **else** Region = character.
End

The proposed rules give an efficient tool for filtering a text object from non-text. They can extract all types of texts from scene images, regardless of the differences between the texts in terms of size, font type, orientation, and the proximity to other objects.

Text Region Filtering using SWT. In this stage, SWT method proposed by Chen, et al. (2011) was employed. The SWT algorithm is based on the idea that stroke is central in text, be it handwritten or typed. Stroke width typically refers to the length of a perpendicular line that connects a text edge pixel to another pixel on the opposite text edge (distance between two pixels on the two parallel text edges). The stroke width is almost consistent throughout a single character in contrast to not-text regions where there is significant change in the stroke width due to their irregularities. In SWT, the first step is to obtain the skeletons of the binary image. For each foreground pixel in the skeleton, distance transform is applied to compute the Manhatten distance from the pixel to the nearest its boundary/edge. This results in a skeleton-distance map. The standard deviation (STD) is then calculated on the skeleton-distance map of each CC strokes to compute the difference between the true text regions and false positives. It should be noted that text characters typically have a much smaller STD compared to the false positive which is fixed. Based on this property, CCs with large STD are removed.

The algorithm of SWT method using Manhattan distance transform for text filtering is expressed in the following equation:

$$\mathrm{M}_{\mathrm{distance}} = |(x_1 - x_2)^2 + |(y_1 - y_2)^2| \tag{5}$$

Algorithm 3: The Proposed Stroke Width Transform for Character Filtering

---

Input: Child Binary Image of a group connected components
Output: Character Region
Start

Step 1: Calculate Manhattan distance transform ($\mathrm{M}_{\mathrm{distance}}$)
Step 2: Calculate stroke width using ($\mathrm{M}_{\mathrm{distance}}$).
Step 3: For each proposed connected component label (CC), calculate the local standard deviation (σ_{stroke}) for each region strokes.
Step 4: Filter the Character by the following threshold equation:

$$T_{extract}\,(x,y) = \begin{cases} \text{black}, \ \sigma_{stroke}) > 1 \\ \textit{white, otherwise} \end{cases} \tag{6}$$

End

Figure 5 illustrates an example result of the proposed SWT method which is used to filter all non-character candidates that were not eliminated in the first filtering step.

RESULTS AND DISCUSSION

The proposed method is evaluated in two ways: visual and analytical. The visual evaluation aims to demonstrate the effectiveness of the proposed method by presenting the visual results for various types of text detection and extraction challenges whilst the analytical evaluation is based on one or more of the benchmark metrics adopted since DIBCO 2009 [Precision and recall, Picture Signal-to-Noise Ratio (PSNR) and Negative Rate Metric (NRM)] to measure the robustness of the proposed method. Two benchmark datasets were used in the evaluations. Firstly, the visual test evaluated the text extraction results by comparing against different state-of-the-art methods. Secondly, the analytical test evaluated the performance of the proposed framework in terms of recognizing the characters and texts (Optical Character Recognition, OCR).

Visual Experiment

The ICDAR Reading Competition organized by the International Conference on Document Analysis and Recognition (ICDAR), has been held five times, in 2003, 2005, 2011, 2013 and 2015. The dataset used in this text reading competition contains about 233 images with ground truth. The KAIST scene text dataset consists of 3000 images captured in various environments, which includes both indoor and outdoor scenes under varying lighting conditions (clear day, night, strong artificial lights). The KAIST dataset is used for evaluation purposes to ensure the coverage of scene text images with different environments and lighting conditions. In addition to the ICDAR and KAIST datasets used for evaluation in this study, we randomly chose several images from another dataset of poor quality such as a degraded document image and a noisy image of license plate. These images are used in the visual assessment for demonstrating the effectiveness of the proposed method in extracting text. Figures 5 to 7 highlights some examples of text extraction results from images that are very difficult to extract.

Figure 8 presents a sample results of the recognition process. A visual analysis of the results demonstrates the effectiveness of proposed OCR framework in accurately recognizing the text particularly from some difficult scene text images containing multi-oriented texts, which the other state-of-the-art methods such as Neumann and Matas (2015) and Gomez and Karatzas (2013) were not able to recognize. The improved results are due to the fact that the proposed text detection and extraction method includes a pre-processing step to prepare and correct difficult images, enhancing it for recognition.

Analytical Experiment

To evaluate the proposed extraction and recognition methods, a comparative study with different state-of-the-art methods in text extraction and recognition was conducted, including the current top ranked methods in ICDAR competition such as Bai et al. (2013), Gomez and Karatzas (2013), Kumar and Lee (2010), Text Spotter by Neumann and Matas (2015), Shi et al. (2013), Sung et al. (2015), Novikova et al. (2012) and Yin et al. (2014). The proposed extraction and recognition methods were evaluated separately. The precision and recall metrics were computed to evaluate the performance of the extraction method whereas accuracy metric was used to evaluate the recognition method. Evaluation was carried out using two benchmark datasets: ICDAR Robust Reading 2013 dataset and KAIST scene image dataset.

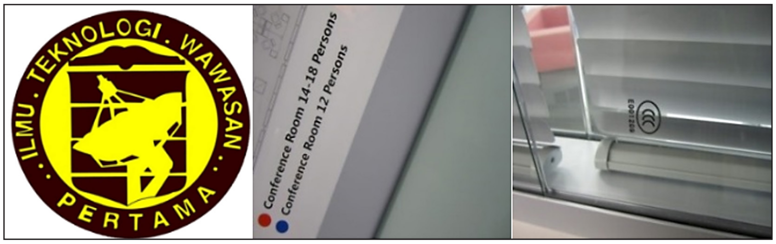
Precision and recall evaluation measures are commonly applied in the area of text extraction (Yin et al., 2014; Neumann & Matas, 2015). Precision is described as the number of text correctly extracted divided by the total number of extracted text and recall is described as the number of text correctly estimated divided by the total number of ground truths (the words found in the original scene image). The precision and recall metrics for text extraction are computed as follows:

$$Precision = \frac{N_{tp}}{N_{tp} + N_{fp}}, \; Recall = \frac{N_{tp}}{N_{tp} + N_{fn}} \qquad (7)$$

where N_{tp}, N_{fp}, and N_{fn}, denote the number of True Positive, False Positive, and False Negative values, respectively.

(a) (b)

Figure 5. Examples of results of the proposed method for images classified as very complex for extraction: (a) original image set on the left side and (b) text extraction results on the right. (*Red and green line border indicates non-text and text regions subsequently)



(a)

(b)

Figure 6. Some example results for non-horizontal text scene image: (a) Original image set and (b) text extraction results

(a)

(b)

Figure 7. Examples of challenges in ICDAR 2013 images and its results using the proposed text detection and extraction methods: (a) original image set and (b) extraction results

(a) (b)

Figure 8. Sample output of the recognition process using the proposed method for more complex images: leftmost image is the original image; center image is the extracted text and rightmost image is the recognition result. (a) Slanted text art shape with complex background, and (b) art text shape from KAIST/ICDAR datasets and google image search engine.

The aim of performance assessment for a text detection and extraction method is to measure the difference between the expected text result and the real text result of the method. Table 4 shows the results for the proposed text detection and extraction methods, comparing it with Sung et al. (2015), Yin et al. (2014), Bai et al. (2013) and Shi et al. (2013) on the ICDAR 2013 dataset. The proposed text detection and extraction methods achieved the best results in both recall and precision measures. The results are expressed as (recall, precision) value pair.: The proposed method obtained (81.04, 88.7) as compared to (72.01, 87.64); (66.45, 88.47); (64.84, 87.51); (68.24, 78.89) for Sung et al. (2015), Yin et al. (2014), Bai et al. (2013) and Shi et al. (2013) respectively.

Table 4
Multi-orientation extraction - evaluation results of the proposed text detection and extraction methods for ICDAR robust reading (2013) and KAIST datasets

Dataset	**ICDAR 2013**		**KAIST**	
Method	**Recall**	**Precision**	**Recall**	**Precision**
Sung et al. (2015)	72.01%	87.64%	-	-
Yin et al. (2014)	66.45%	88.47%	-	-
Bai et al. (2013)	64.84%	87.51%	-	-
Shi et al. (2013)	68.24%	78.89%	78%	66%
Kumar and Lee (2010)	-	-	60%	69%
Proposed method	81.04%	88.71%	83%	81%

As shown in Table 4, the proposed technique yielded a higher recall and precision values than any other state-of-the-art methods using the ICDAR dataset. Figures 5 and 6, which illustrate a sample of the extraction results on horizontal texts and non-horizontal text, respectively, demonstrate the effectiveness of the proposed method in detecting many text objects, including text with different colors, different lightings, complex backgrounds, different sizes, different stroke widths, flexible surfaces, other symbols, non-horizontal directions as well as low contrasts.

For the evaluation using the KAIST dataset, the proposed detection and extraction method outperformed Shi et al. (2013) and Kumar and Lee (2010) in extracting text from more complex and challenging scene images such as various font sizes, colors and abrupt environmental changes. The results in Table 4 shows the proposed detection and extraction methods achieved its best performance in both precision and recall whereby the proposed method achieved (0.81, 0.83) as compared to (0.66, 0.78) and (0.69, 0.60) for Shi et al. (2013) and Kumar and Lee (2010) respectively. From these results, it is observed that the proposed detection and extraction methods were able to extract text from the scene image with minimum negative error for different text width sizes and variants of text color. Furthermore, the proposed detection and extraction method also solved the problem of extracting characters for the exclusive case such as text art case outperforming other state-of-the-art methods (Figure 8). In summary, these results have been evaluated according to the precision and recall metrics with respect to the proportions of the contents of the texts that have been extracted in comparison to the ground truth.

Despite the achievement, the strength and shortcoming of the state-of-the-art approaches are discussed below:

1. Sung et al. (2015) introduced a three-stage framework for character extraction, character verification and character refinement phases. Their character extraction mechanism consisted of ER tree construction, sub-path partitioning, sub-path pruning, and character candidate selection sequentially. Apart from employing AdaBoost trained character classifier for character verification, they commissioned heuristic rules for character and group refinement prior to geometric adjacency and color analysis. Regardless of its complexity and abandoning low contrast issue, the three-stage framework indicated significant results when dealing with multi-color and multi-oriented text.
2. Maximally Stable Extremal Regions (MSERs) algorithm initiated by Yin et al. (2014) was able to prune character candidates with the strategy of minimizing regularized variations. Furthermore, their single-link clustering algorithm so called a novel self-training distance metric learning algorithm was able to distinguish text and non-text areas automatically and efficiently. This innovation can be further accelerated if non-uniform illumination strategy is accounted.
3. Bai et al. (2014) proposed gradient local correlation for efficient scene text localization. In compliance to noise, small character size and shadow insensitivity, they developed a text confidence map by characterizing the density of pair wise edges and stroke width consistency and segmented text and non-text area using simple connected component, SVM classifier and color analysis.
4. By utilizing the available contrast of text pixels, Kumar and Lee (2010) claimed that their text extraction method was able to recover the missing text area. It comprised a set of algorithms such as pixel quantization, color to grey conversion, geometrical features, three layers of connected component and Sobel Projection which was able to recover missing text from the non-text region. Despite of its outstanding performance in dealing with non-uniform illumination, this handcrafted method could be further explored for achieving better results in non-uniform and oriented text.

For the text recognition evaluation using ICDAR 2013 and KAIST datasets, accuracy measure (Equation 8) which is the ratio of the number of correctly recognized words to the total number of words in the dataset is used. Accuracy is used instead of precision as the intention is just to demonstrate the performance of the proposed OCR method in recognizing the words in the dataset compare to the existing methods.

$$Accuracy = \frac{N_{tp} + N_{tn}}{N_{tp} + N_{fp} + N_{tp} + N_{tp}} \quad (8)$$

where N_{tp}, N_{fp}, N_{tn}, and N_{fn}, denote the number of True Positive, False Positive, True Negative and False Negative values, respectively. The evaluation results are tabulated in Table 5.

The results of the text recognition evaluation using ICDAR 2013 dataset in Table 5 show that the proposed OCR framework outperformed Novikova et al. (2012) method and was on par with Shi et al. (2013) method in terms of the number of words correctly recognized where the proposed method achieved 81.23% when compared to 82.1% and 57.99% for Shi et al. (2013) and Novikova et al. (2012) methods, respectively. Some sample results are highlighted in Figure 7.

For the KAIST dataset, results in Table 5 show that the proposed framework yielded better result than other methods in the recognition phase with the OCR accuracy of approximately 68% comparing to 56% and 63% for Text Spotter by Neumann and Matas (2015) and Gomez and Karatzas (2013), respectively. Furthermore, KAIST dataset contains categories of scene images according to the types of environments (outdoor, light and night, indoor and shadow). Table 6 demonstrates the effectiveness of the proposed OCR framework in the different types of environment. The proposed OCR method achieved high rate recognition in the outdoor and shadow environment types, 87% and 86.01% respectively.

Table 5
Text recognition evaluation results for ICDAR robust reading (2013) and KAIST datasets

Dataset	**Method**	**Percentage of correctly recognized words**
ICDAR Reading	Shi et al. (2013)	82.1%
	Novikova et al. (2012)	57.99%
	Proposed method	81.23%
KAIST	Text Spotter Neumann and Matas (2015)	56.4%
	Gomez and Karatzas (2013)	62.99%
	Proposed method	68.06%

Table 6
Results of the proposed OCR framework in KAIST dataset pertaining to four environments

Type	**No of words**	**Recognized words**	**OCR accuracy**
Outdoor	184	160	87.0
Light & night	38	27	71.1
Indoor	91	62	68.1
Shadow	36	31	86.01

The following are some benefits and limitations of the state-of-the-art word recognizers that may affect the experimental results:

Novikova et al. (2012) integrated local likelihood and pairwise positional consistency priors mainly for enforcing consistency of characters (lexicon) and their attributes in terms of font and color. Their word recognition process successfully estimated the maximum a posteriori or MAP inference under the joint posterior distribution of the model namely weighted finite-state transducers. Perhaps the mapping of font and color attributes in MAP inference requires further improvising.

Shi et al. (2013) applied part-based tree-structure to detect and recognize each type of character simultaneously. Their framework included modelling potential character locations using Conditional Random Field that incorporated detection scores, spatial constraints and linguistic knowledge. In the final stage, their word recognition result was achieved through the usage of cost function minimization in the random field.

Gomez and Karatzas (2013) also initiated a perceptual framework that exploited collaboration of proximity and similarity laws to create text-group hypotheses.

Driven by similar motivation, Neumann and Matas (2015) introduced a two-stage approach comprises a sequential selection against time from the set of Extremal Regions (ER) and clustering algorithm. They claimed ER was robust against blur, low contrast and illumination, color and texture variation because only ERs with locally maximal probability were nominated for the classification phase.

Overall, our proposed method has been shown to be effective in recognizing text of different forms i.e. size, color, stroke width, and orientation under different environmental condition i.e. outdoor, indoor, light and night, and shadow, outperforming the benchmarked state-of-the-art methods. Our method works especially well under the outdoor and shadow environments. Further work needs to be carried out to improve text recognition in the light and night as well as indoor categories. A robust text recognition method that is reasonably accurate in recognizing different fonts across a variety of environment is particularly attractive for applications that involve mobility such as detecting car license plate on vehicles on the highways as well as in indoor car park areas, providing scene information to support the visually impaired people via scene-to-text-to-speech translation as they carry out their daily activities in various environments and recognition.

CONCLUSION

In this paper, a novel text detection/extraction and recognition technique for real scene text images that are tolerant of different types of degradation is presented. The main objective of the proposed detection and extraction method and OCR framework is to tackle the extraction issues and difficulties faced when dealing with natural scene text images, specifically in solving multi-oriented text, the lighting effect on the appearance of text, and low contrast. The proposed text extraction and detection method consists of three steps: first is the exploited NTSC enhancement, second is the hybrid adaptive segmentation combining fuzzy C-means segmentation and adaptive binarization, and third is a rule-based text extraction with stroke width transform using statistical and geometrical features during text filtering for detecting connected component regions. After distinguishing text from non-text extracted from a scene image, text recognition is accomplished using the Tesseract OCR engine.

To evaluate the proposed method, it was compared with the state-of-the-art detection/extraction and recognition methods. All the experiments were conducted using the ICDAR 2013 robust reading dataset and KAIST scene text dataset. Experimentally, it is established that the proposed method outperformed current state-of-the-art extraction methods, achieving >80% in both precision and recall measures. This is because the proposed method enhances the quality of the scene text image source by solving most of the text detection problems such as high-orientation text, low contrast, and illumination for better text extraction and recognition.

ACKNOWLEDGEMENTS

The authors would like to thank the Faculty of Information Science and Technology and the Center for Artificial Intelligence Technology, Universiti Kebangsaan Malaysia for providing the facilities and financial support under the grant: AP2017-05/2 and FRGS-1-2019-ICT02-UKM-02-9.

REFERENCES

Bai, B., Yin, F., & Liu, C. L. (2013, August 25-28). Scene text localization using gradient local correlation. In *2013 12th International Conference on Document Analysis and Recognition (ICDAR)* (pp. 1380-1384). Washington, DC, USA.

Behzadi, M., & Safabakhsh, R. (2018, December 25-27). Text Detection in Natural Scenes using Fully Convolutional DenseNets. In *2018 4th Iranian Conference on Signal Processing and Intelligent Systems (ICSPIS)* (pp. 11-14). Tehran, Iran.

Burney, S. A., & Tariq, H. (2014). K-means cluster analysis for image segmentation. *International Journal of Computer Applications*, *96*(4), 1-8.

Chen, H., Tsai, S. S., Schroth, G., Chen, D. M., Grzeszczuk, R., & Girod, B. (2011, September 11-14). Robust text detection in natural images with edge-enhanced maximally stable extremal regions. In *2011 18th IEEE International Conference on Image Processing (ICIP),* (pp. 2609-2612). Brussels, Belgium.

Gomez, L., & Karatzas, D. (2013, August 25-28). Multi-script text extraction from natural scenes. In *2013 12th International Conference on Document Analysis and Recognition* (pp. 467-471). Washington, DC, USA.

Horvath, J. (2006, January 20-21). Image segmentation using fuzzy c-means. In *Proceedings of SAMI, 4th Slovakian-Hungarian Joint Symposium on Applied Machine Intelligence* (pp. 144-151). Herlany, Slovakia.

Huang, S., Xu, H., Xia, X., & Zhang, Y. (2018, December 8-9). End-to-end vessel plate number detection and recognition using deep convolutional neural networks and LSTMs. In *2018 11th International Symposium on Computational Intelligence and Design (ISCID)* (Vol. 1, pp. 195-199). Hangzhou, China.

Jadhav, A. J., Kolhe, V., & Peshwe, S. (2013). Text Extraction from Images: a survey. *International Journal of Advanced Research in Computer Science and Software Engineering*, *3*(3), 333-337.

Jung, K., Kim, K. I., & Jain, A. K. (2004). Text information extraction in images and video: a survey. *Pattern Recognition*, *37*(5), 977-997.

Karanje, U. B., & Dagade, R. (2014). Survey on Text Detection, Segmentation and Recognition from a Natural Scene Images. *International Journal of Computer Applications*, *108*(13), 39-43.

Kraisin, S., & Kaothanthong, N. (2018, November 15-17). Accuracy Improvement of a Province Name Recognition on Thai License Plate. In *2018 International Joint Symposium on Artificial Intelligence and Natural Language Processing (iSAI-NLP)* (pp. 1-6). Pattaya, Thailand.

Kumar, M., & Lee, G. (2010, February 26-28). Automatic text location from complex natural scene images. In *2010 The 2nd International Conference on Computer and Automation Engineering (ICCAE)* (Vol. 3, pp. 594-597). Singapore.

Kumar, M., Kim, Y. C., & Lee, G. S. (2010, June 29-July 1). Text detection using multilayer separation in real scene images. In *2010 IEEE 10th International Conference on Computer and Information Technology (CIT)* (pp. 1413-1417). Bradford, UK.

Li, H., Wang, P., & Shen, C. (2018). Toward end-to-end car license plate detection and recognition with deep neural networks. *IEEE Transactions on Intelligent Transportation Systems, 20*(3), 1126-1136.

Nagaraju, G., Ramaraju, P. V., Sandeep, P. M. M. P., Nawaz, S. M., & Bhargav, S. K. (2015). Text Extraction from Images with Edge-Enhanced Mser and Hardware Interfacing Using Arduino. *International Journal of Engineering and Computer Science*, *4*(05), 11797-11803.

Neumann, L., & Matas, J. (2015). Real-time lexicon-free scene text localization and recognition. *IEEE transactions on pattern analysis and machine intelligence*, *38*(9), 1872-1885.

Ning, G., Han, T. X., & He, Z. (2015, September 27-30). Scene text detection based on component-level fusion and region-level verification. In *2015 IEEE International Conference on Image Processing (ICIP)* (pp. 837-841). Quebec City, QC, Canada.

Nosal, E. M. (2008, October 14-17). Flood-fill algorithms used for passive acoustic detection and tracking. In *2008 New Trends for Environmental Monitoring Using Passive Systems* (pp. 1-5). Hyeres, French.

Novikova, T., Barinova, O., Kohli, P., & Lempitsky, V. (2012). Large-lexicon attribute-consistent text recognition in natural images. In *European Conference on Computer Vision* (pp. 752-765). Heidelberg, Germany: Springer.

Phan, T. Q., Shivakumara, P., & Tan, C. L. (2009, July 26-29). A Laplacian method for video text detection. In *2009 10th International Conference on Document Analysis and Recognition* (pp. 66-70). Barcelona, Spain.

Qaisar, S. M., Khan, R., & Hammad, N. (2019, March 26-April 10). Scene to Text Conversion and Pronunciation for Visually Impaired People. In *2019 Advances in Science and Engineering Technology International Conferences (ASET)* (pp. 1-4). Dubai, United Arab Emirates.

Raghunandan, K. S., Shivakumara, P., Roy, S., Kumar, G. H., Pal, U., & Lu, T. (2018). Multi-script-oriented text detection and recognition in video/scene/born digital images. *IEEE Transactions on Circuits and Systems for Video Technology, 29*(4), 1145-1162.

Raj, H., & Ghosh, R. (2014, September 24-27). Devanagari text extraction from natural scene images. In *2014 International Conference on Advances in Computing, Communications and Informatics (ICACCI)* (pp. 513-517). New Delhi, India.

Rajaby, E., Ahadi, S. M., & Aghaeinia, H. (2016). Robust color image segmentation using fuzzy c-means with weighted hue and intensity. *Digital Signal Processing*, *51*, 170-183.

Risnumawan, A., Shivakumara, P., Chan, C. S., & Tan, C. L. (2014). A robust arbitrary text detection system for natural scene images. *Expert Systems with Applications*, *41*(18), 8027-8048.

Samadhiya, A., & Khatri, P. (2014). General Review of Various Text. *International Journal of Advances in Electronics and Computer Science, 2*(1), 12-16.

Shah, K. M. & Thakar, V. K. (2010). Content Based Image Retrieval using different Colormaps. In C. Modi, H. Shah, R. Kher & N. Desai (Eds.), *Proceedings of the 2009 International Conference on Signals, Systems and Automation (ICSSA 2009)*. Boca Raton, Florida: Universal-Publishers.

Shi, C., Wang, C., Xiao, B., Zhang, Y., Gao, S., & Zhang, Z. (2013, June 23-28). Scene text recognition using part-based tree-structured character detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 2961-2968). Portland, Oregon, USA.

Shivakumara, P., Phan, T. Q., & Tan, C. L. (2011). A laplacian approach to multi-oriented text detection in video. *IEEE transactions on pattern analysis and machine intelligence*, *33*(2), 412-419.

Sulaiman, A., Omar, K., & Nasrudin, M. F. (2019). Degraded historical document binarization: A review on issues, challenges, techniques, and future directions. *Journal of Imaging, 5*(4), 48-73. doi: 10.3390/jimaging5040048

Sumathi, C. P., Santhanam, T., & Devi, G. G. (2012). A survey on various approaches of text extraction in images. *International Journal of Computer Science and Engineering Survey*, *3*(4), 27-42.

Sun, Y., Mao, X., Hong, S., Xu, W., & Gui, G. (2019). Template Matching-Based Method for Intelligent Invoice Information Identification. *IEEE Access, 7*, 28392-28401.

Sung, M. C., Jun, B., Cho, H., & Kim, D. (2015, August 23-26). Scene text detection with robust character candidate extraction method. In *2015 13th International Conference on Document Analysis and Recognition (ICDAR)* (pp. 426-430). Tunis, Tunisia.

Wang, L., Huang, L. L., & Wu, Y. (2011, November 28). An efficient coarse-to-fine scheme for text detection in videos. In *2011 First Asian Conference on Pattern Recognition (ACPR)* (pp. 475-479). Beijing, China.

Wang, Q., Lu, Y., & Sun, S. (2015, August 23-26). Text detection in nature scene images using two-stage nontext filtering. In *2015 13th International Conference on Document Analysis and Recognition (ICDAR)* (pp. 106-110). Tunis, Tunisia.

Yao, J. L., Wang, Y. Q., Weng, L. B., & Yang, Y. P. (2007, November 2-4). Locating text based on connected component and SVM. In *2007International Conference on Wavelet Analysis and Pattern Recognition (ICWAPR'07)* (Vol. 3, pp. 1418-1423). Beijing, China.

Yin, X. C., Yin, X., Huang, K., & Hao, H. W. (2014). Robust text detection in natural scene images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *36*(5), 970-983.

Zhang, H., Zhao, K., Song, Y. Z., & Guo, J. (2013). Text extraction from natural scene image: A survey. *Neurocomputing*, *122*, 310-323.

Pertanika J. Sci. & Technol. 27 (4): 1783 - 1802 (2019)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Time Adaptive Collaborative Filtering for Movie Recommendation

Suganeshwari Gopalswamy and Syed Ibrahim Peer Mohamed*

School of Computing Science and Engineering, Vellore Institute of Technology, Chennai, Tamil Nadu 600 127, India

ABSTRACT

Collaborative filtering is the most widespread recommendation system technique deployed in e-commerce services nowadays. It recommends products based on the historical preference of the user. The biggest challenges in these techniques are data sparsity and growing volume of data, specifically in e-commerce sites like movie recommendation. Clustering algorithms are used for scaling up the performance of collaborative filtering in dynamically growing datasets. Most of the existing clustering based recommendation algorithms improve scalability but produce low quality recommendations. This is mainly due to data sparsity, as the user tends to rate very few items from a large number of options available. Moreover, users with a similar taste for a group of items may show different likings for another group of items over a period, i.e.., user's interest dynamically changes over time. Finding the sub-groups that are more relevant to each other than the entire user-item matrix is more affordable. Since the user's recent ratings can better represent their interest and preference, a Time Adaptive Collaborative Filtering Method —TACF is proposed, that adopts time to generate a recommendation. Experimental results on the MovieLens dataset show that the proposed system outperforms other state-of-art collaborative filtering algorithms in terms of accuracy and efficiency.

Keywords: Clustering, collaborative filtering, matrix factorization, recommendation system, temporal information, user drifts

INTRODUCTION

In recent years, the recommender system has emerged as a tool to cope up with the information overload problem. With the growing volume of information on the web, it has become inevitable for companies to provide the most relevant information to the user. Currently, companies are deploying

ARTICLE INFO

Article history:
Received: 24 January 2019
Accepted: 04 Jun 2019
Published: 21 October 2019

E-mail addresses:
g.suganeshwari2015@vit.ac.in (Suganeshwari Gopalswamy)
syedibrahim.sp@vit.ac.in (Syed Ibrahim Peer Mohamed)
* Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

intelligent recommender systems to personalize recommendations according to the customer's taste. The Recommendation system has been a great success in dealing with an information overload problem generating more relevant recommendation in a variety of applications like recommending movies, CDs, books and webpages.

Recommendation System (RS) is broadly classified as content-based, collaborative filtering (CF) and hybrid approaches. Content-based approaches generate a user profile based on the purchase history of the user (Desrosiers & Karypis, 2011). This profile is compared with item features and then suitable recommendations are generated. Collaborative filtering (Sarwar et al., 2001) provides a recommendation based on the idea of people who share a similar taste in certain items sharing the same interest in the rest of the items. It works well for complex items like movies (Harper & Konstan, 2016) music (Tan et al., 2011) and tourism (Jiang et al., 2016). Hybrid methods work by combining the content and collaborative approaches.

CF method is the most popular method used in the recommender system. It is further classified as memory and model-based approaches (Adomavicius & Tuzhilin, 2005). It finds extensive implementation in the e-commerce industry due to its justifiability and ease of implementation (Desrosiers & Karypis, 2011). However, this method suffers from data sparsity, which occurs mainly as a user's rate only a few products from a large number of available options. Furthermore, when the number of users and items increase, the computational complexity also increases leading to poor scalability becoming infeasible to produce recommendations within the elapsed time interval.

Model-based techniques have been introduced for meeting the challenges from memory-based CF. These techniques discover the rating pattern from historical data and provide highly accurate and efficient recommendations from some sample data. Bayesian network (Luo et al., 2012), Clustering model (Li & Kim, 2003) (West et al., 2016) and matrix factorization (Takács et al., 2008) are familiar techniques used in model-based approaches. SVD, SVD++ (Koren, 2009) and ALS (Chen et al., 2017) are popular MF techniques used in the recommender system that gained importance following the Netflix price challenge (Bell & Koren, 2007).

In the clustering-based recommendation approach, the users and items are segmented in such a way that objects in the same cluster have a large similarity. The CF method gathers ratings from all the users and then computes the predictions. It is a time-consuming process and in the previous work (Suganeshwari & Ibrahim, 2016) has focused on the incorporation of the temporal information for generating a time-aware recommendation and improving efficiency. However, despite the improved performance in recommendation quality, scalability still remains an issue in a real-time environment. Motivated by these observations and understanding of the likelihood of popular items causing bias in the recommendations generated, we propose a Time Adaptive Collaborative Filtering (TACF).

This paper provides a scalable algorithm which shows a significant improvement in precision on recommending top-k items when compared with other existing algorithms ALS, UBCF, IBCF, and time-sensitive CF (TSCF) (Sun et al., 2016). The experiments have been performed using benchmark MovieLens dataset.

The rest of the paper is organized as follows. Section II provides an introduction to the preliminary concepts of RS and discusses its related work and its limitations. Section III presents the proposed method Time Adaptive Scalable Neighborhood (TACF). Section IV describes the experimental setup and enlists the evaluation results followed by the conclusion.

BACKGROUND AND RELATED WORK

Section 2 introduces the preliminaries on memory-based and model-based collaborative filtering methods. Table 1 represents the symbol and their representation used in this paper.

Table 1
Table of symbols

Symbol	Definition
U	Set of users
I	Set of items
R^{mxn}	Utility matrix
u	User
i	Item
u_{ij}	User i rating on item j
W	Item clusters
K	Dimension of the user and item latent factors
P	User latent factor matrix
Q	Item latent factor matrix
λ	Regularization parameter
α	# of neighbors
δ	Recent transactions of user
S_{ij}	Similarity between item i and j
b_u	User bias
b_i	Item bias

Memory-Based Collaborative Filtering

Memory-Based CF methods are also referred to as neighborhood methods. They recommend items on the basis of the views of other like-minded people. Similarity computation plays a vital role in CF methods (Herlocker et al., 2002). The most popular statistical techniques used for computing similarity are Pearson correlation (Adomavicius & Tuzhilin, 2005) and cosine based (Meng et al., 2014) measures. Here the entire utility matrix of size *User X Item* is taken as input. The similarity is evaluated by employing user correlation in user-based (Meng et al., 2014) and item correlation in item based (Sarwat et al., 2014) methods. The bottleneck in CF methods computation of the similarity between users/items. Though these methods pioneer the e-commerce industry, it is affected by data Sparsity. The neighborhood-based CF methods rely on exact matches for computing correlations resulting in an extreme sparse dataset. This insufficient *user X item* matrix leads to inaccurate predictions where similar items do not show any correlation and are termed as reduced coverage (Billsus & Pazzani, 1998). Traditional CF algorithms suffer in scalability when dealing with the tremendously growing volume of data. Several methods have been proposed for dealing with problems like imputation techniques (Ren et al., 2013), and incorporation of the contextual parameters that include location, the company of other people and time. LARS, a location-aware recommendation system has been proposed that provides recommendations on the basis of spatial ratings (Sarwat et al., 2014). Time is used as a special type of context in Ding & Li, (2005) and Panniello et al. (2014) for improving the recommendation quality. Memory-based methods rely on the similarity computation between users/items for the determination of the predictions, but changes in user preference drifts can mislead recommendations that may not be of interest to the user in the current situation. Time is used to compute the similarity between the users in a Time-aware RS (Hu et al., 2015). Ratings within the similar timestamps are assigned a high weight than the other older ratings by employing an exponential decay function. The function value drops rapidly when the time difference increases. It is a time-weighted approach where the older values are penalized. This method cannot be deployed in movie recommendation application as user rates the movie only once and the rating remains static. Sun et al., (2016) proposed a time-sensitive CF (TSCF) approach to discover the latest preference of the user by ordering the items based on the time behavior sequence. The system considers only one transaction which is insufficient to generate better recommendations. The proposed method claims that recent *n* transaction of each user is adequate to compute similarity and to reflect the user's current preference.

Model-Based CF

In real-time applications deployment of the memory-based method is infeasible especially with the explosive increase in the number of items and users. The model-based CF utilizes the rating matrix to construct a model and later exploits it for future predictions, thus improving the efficiency of the system. Different machine learning techniques like classification, clustering, and matrix factorization are employed to construct the model. The classification algorithms used in CF are Simple Bayesian Algorithm (Miyahara & Pazzani, 2000 and Su & Khoshgoftaar, 2006) and Baseline Bayesian model (Heckerman et al., 2000). Matrix Factorization is a widely used dimensionality reduction technique in RS. An extensive survey on recommendation (Takács et al., 2008 and Ba et al., 2013) shows the help rendered by the application of matrix factorization techniques in revealing the latent factor. The results can be easily interpreted for predicting the rating for an item.

User's interest is always drifting with time and a dynamic model based RS is needed to address this. The user rating behavior is influenced by the user rating style and the item's popularity. SVD++ incorporated temporal information to improve the quality of recommendation (Bell & Koren, 2007). The lessons learnt from Netflix price challenge (Bell & Koren, 2007), help the neighborhood method and model-based methods in the exploration of different levels of data patterns and hybrid methods alone have the ability to produce more optimal results. In the proposed work, a time-adaptive hybrid model is designed to address the scalability and sparsity problem.

Large-Scale Datasets and Clustering

Data scalability is one of the key challenges in providing recommendation in a real-time environment. This occurs mainly as a result of the tremendous growth of users and items in modern e-services. The aim of clustering algorithms is to partition the data into meaningful subgroups. High-quality clusters are produced when elements within the cluster are more similar and elements from different clusters are dissimilar. Clustering algorithm segments items based on the user rating data (O'Connor & Herlocker, 1999). Similarly, partition based on user-user similarity was proposed by Sarwar & Karypis (2002a). Xue et al. (2005) used the clustering method to smooth the missing value in the utility matrix. West et al. (2016) had proposed a scientific article based on the citation-based network. It used a hierarchical representation of scientific structure as domains, fields, subfields, and sub-subfields for different levels of influence. Scalability and sparsity problems were addressed in Ma et al. (2016) where different clustering techniques were adopted on the basis of the user, item and trust relationships. Guo et al. (2015) had utilized the rating information along with social trust information for iterative segmentation of users. There are myriad applications in a recommendation system that exploits the clustering methods. All these methods get adversely affected due to sparsity, resulting in a low-quality recommendation. Hence, we argue that better recommendations can be produced if the data to be clustered is dense. The proposed TACF method applies a clustering algorithm to produce clusters based on latent factors and exploits time to generate recommendations. The proposed method can improve the efficiency and mean average precision of the CF method, and this is demonstrated in the experimental results.

TIME ADAPTIVE COLLABORATIVE FILTERING FRAMEWORK

The primary objective of TACF is to find the item sub-groups drowned in the *user x item* utility matrix for improving the quality of the CF recommendation algorithm. It is difficult to identify the retention of the user's interest in the item as the user rates an item only once. To address this problem, the matrix factorization method is utilized where the user and item features are found and used as input to the clustering method. The features are more compact and denser. Items are clustered based on item features. Recommendations are generated based time adaptive function. The proposed method fully considers the user's preference drifts, addresses sparsity issues and guarantees the achievement of better performance.

Problem Formulation

With the implication of the presence of n user's $U = \{u_1, u_{2, \ldots,} u_n\}$ and m items $I = \{i_1, i_{2\ldots,} i_m\}$ represented as matrix R of size m x n. Each user's rating is represented as u_{ij} where u_i represents the i^{th} user and u_j represents the j^{th} item. The aim is to divide the items into w subgroups $\{w_1, w_2\ldots, w_c\}$. This is the clustering problem, but a matrix factorization model comprising user and item factors has been created for framing more meaningful segments. The rating matrix is decomposed into lower rank matrices P and Q^T using the ALS method as given in equation 1.

$$R = PQ^T \qquad (1)$$

$P \; \varepsilon \; R^{mxk}$ represent the user latent factors, $Q \; \varepsilon \; R^{kxn}$ represent the item latent factors, k is the latent factor formed and T is the transpose operator. The aim of this function is to frame P and Q^T such that the Frobenius norm is minimized. Once the factor matrices are built, the x^{th} row vector $p_x \; \varepsilon$ P represents the user latent factor and y^{th} column vector $q_y \; \varepsilon$ Q represents the item latent factor In the proposed method, the item factors are used as inputs to the clustering model as they generate subgroups based on the user-rating behavior.

The TACF encodes the item factors Q^T into w clusters and represents them as w ε [0,1]. Here each data point w_{ij} is a value that indicates the presence of the element in the subgroup. Where w_{ij} =1 the element is in the subgroup else, it is not present in the subgroup. The item factor Q^T is formally partitioned into w_c $\{w_1 \; U \; w_2 \ldots w_c\} = Q^T$, w_x □ $w_y = \phi$, for $x \leq 1$ and $y \leq c$ as shown in equation 2.

$$Q^T = \{w_1, \; w_2, \; \ldots \; w_c\} \qquad (2)$$

The squared Euclidean distance is used in classical clustering.

$$K_p = \sum_{i=1}^{n} \sum_{j=1}^{c} u_{i,j}\left(p_{i,t}\right) \qquad (3)$$

Design of TACF

TACF proposes a scalable neighborhood formulation for improving the efficiency of collaborative filtering. The steps involved in the algorithm framework are shown in Figure

1. In the first stage, an ALS model is created from the user-item matrix R. The decomposed matrix consists of user features and item features. This item factor acts as input to the TACF model.

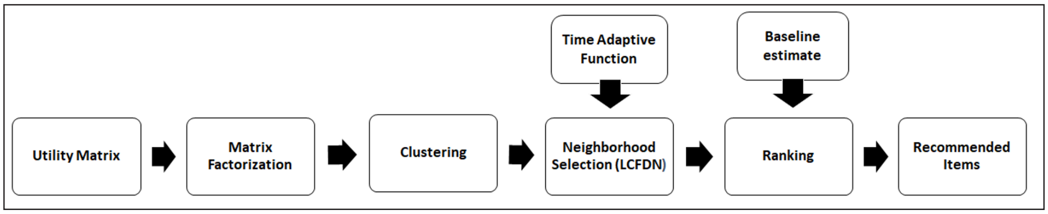


Figure 1. Steps in TACF method

Matrix Factorization

Matrix Factorization is the most popular CF algorithm used for solving the co-clustering problem. CF methods produce accurate recommendations on dimensionality reduced dataset rather than sparse dataset (Sarwar et al., 2001). The utility matrix of dimension U x I get reduced to U x K and I x K. MF techniques can take the input both explicit and implicit. User's rating for an item can be predicted through simple multiplication of the matrix as given in equation 4.

$$r_{u,i} = p_u^T q_i \tag{4}$$

The following quadratic function should be minimized for learning the factors obtained from the factorization techniques.

$$argmin_{p,q} \sum_{u,i} \left(r_{u,i} - p_u^T q_i\right)^2 - \lambda(|p_u|^2 + |q_i|^2) \tag{5}$$

$\sum_{u,i}\left(r_{u,i} - p_u^T q_i\right)^2$ is the Mean square error of the original matrix and the approximation matrix. The $\lambda(|p_u|^2 + |q_i|^2)$ is the regularization term added for avoiding the overfitting of the data. The idea is to produce a minimum value for the cost function using the parameters k and λ. The popular methods used in recommendations are SVD, SGD and ALS.

Singular Value Decomposition (SVD). SVD is the well-established factorization technique used in RS. In SVD, the matrix R gets reduced to three matrices p, σ, and q.

$$r_{u,i} = p\sigma q^T \tag{6}$$

Here p represents the left singular matrix in which each user is represented in some latent factor, q represents the right singular matrix in some latent factor and $\sum$ represents the strength of the latent factor. Maximum Margin Matrix Factorization is a low-rank approximation (DeCoste, 2006) that addresses the noise problem in CF. SVD has been extensively researched in CF methods, but it suffers from computational complexity and poor recommendation when data is sparse.

Stochastic Gradient Descent (SGD). In SGD (Yu et al., 2014) the cost functions are computed, and the factors in the opposite side of the opposite side of the gradient are updated. SGD is not practical when the explicit preferences are converted to implicit ones. User time items would be in the order of huge numbers. SGD requires many numbers of iterations for obtaining the good model. Its performance is also based on the learning rate (Park et al., 2017). This has helped the proposed method in the adoption of the alternative least square method which is more efficient

Alternative Least Square. Alternate Least Square (ALS) is another optimizing technique that can handle data sparsity and still achieve good performance (Chen et al., 2017). The main advantage of this method is its parallel implementation and suitability for large datasets. TACF utilizes explicit user input as ratings. P_u and $Q_i \varepsilon R^f$ are the user and item vectors that represent the value that measures the latent factor, the user and item possessed. By calculating the partial derivative of p_u, q_i and substituting it to 0, equation 7 and equation 8 are obtained.

$$p_u = \left(Q^T Q + \lambda I\right)^{-1} Q^T r_u \tag{7}$$

$$q_i = \left(P^T P + \lambda I\right)^{-1} P^T r_i \tag{8}$$

Here I is the unit matrix, r_u, r_i is the u^{th} row and i^{th} column of the matrix R. The solution given by ALS is unique and the model is constructed until convergence. ALS is adopted in TACF since it has the advantage of parallel implementation with the ability to attain accurate recommendations in the sparse dataset. After decomposing the original matrix R into pq^T, the item factor vector q_i^T is used as input to the clustering method.

Clustering Items

Clustering methods groups objects in segments where members within the cluster have similar features and members of different clusters are dissimilar. This dimensionality reduction technique produces low-quality recommendation when compared to the nearest neighborhood method (Sarwar & Karypis, 2002b). K-means clustering partitions the data points into *k* distinct clusters. The proposed work utilizes the Euclidean distance measure. The number of clusters (K) is selected based on the elbow method and shown in Figure 2. The objective function of the *K-means* is to minimize the squared error function as in equation 9.

$$J = \Sigma_{i=1}^{n} \Sigma_{j=1}^{x} |x_i - c_j|^2 \qquad (9)$$

The utility matrix is decomposed into item and user factors using the ALS method for dealing with these issues. The item factors are then given as input to the clustering method. The clusters framed from the item factors provide a better relationship between the items when compared to the classical K-means clustering. In each round, each point is examined with centroid to find the closest cluster. So, the computational complexity is O(KN). Here k is a number of clusters and n is the number of points. But in a sparse dataset, the number of iterations to converge can be very large. The K-means converges in a fewer number of iterations when the points are represented in a low dimensional matrix using MF as shown in Figure 3.

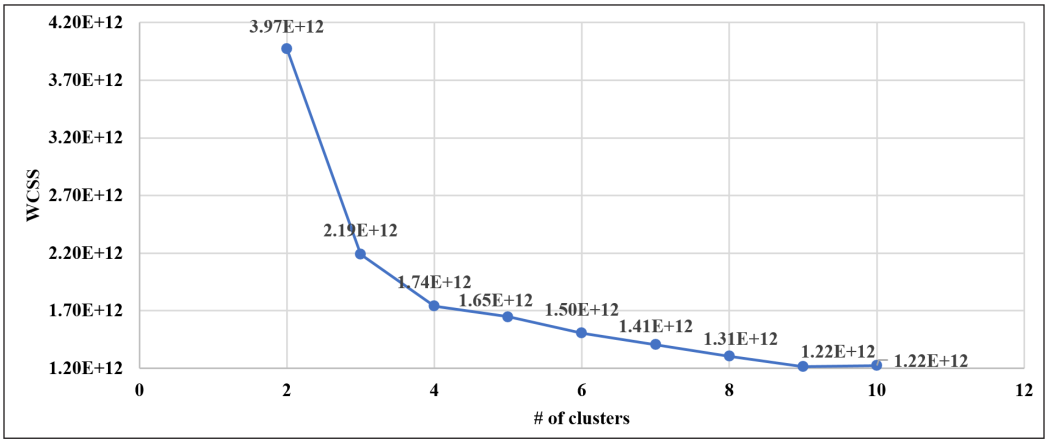


Figure 2. Number of clusters along With in Cluster Sum of Squared Errors (WCSS) in K-means for 1M dataset

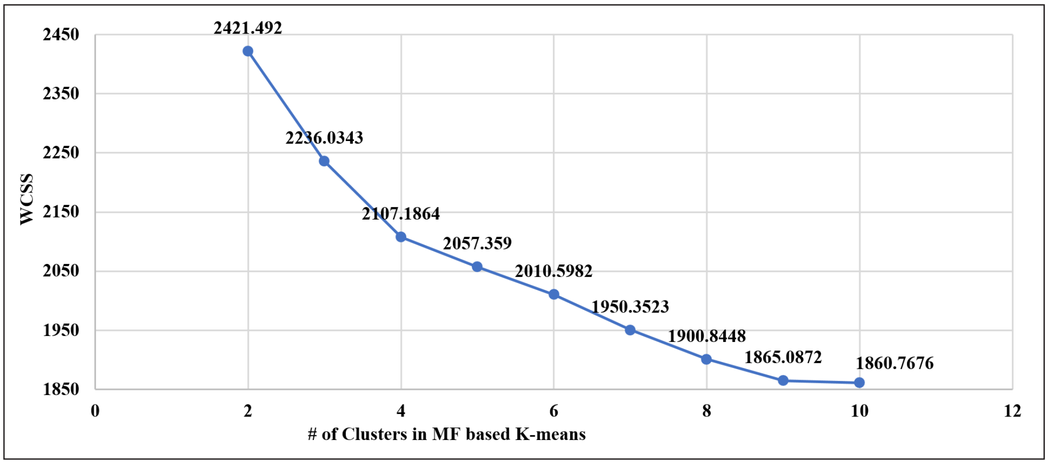


Figure 3. Number of clusters along With in Cluster Sum of Squared Errors (WCSS) in MF based K-means for 1M dataset

Lazy Collaborative Filtering with Dynamic Neighborhood

TACF utilizes the LCFDN (Suganeshwari et al., 2018) approach as its primal recommendation method. The clusters framed are meaningful but do not address the user preference drifts of the user. The LCFDN concentrates on improving the accuracy of RS by adopting time with two defined features α and δ. Recommendations generated based on time yielded better quality and accurate results.

Phase 1: Prediction Computation. A numerical value is formulated by computing the weighted average given in equation 10. Computation of similarity is done between items i and j is done by the isolation of ratings given by the users to both the items and statistical measures is applied. In this work the correlation ratio found is shown in equation 11. Using these similarity ratios for each item i ε I, a similar items list s is stored with k items with highest similarity ratios sim (i, j).

$$P_{u,i} = \frac{\sum_{\delta,\ n}\left(S_{i,n} * R_{u,n}\right.}{\sum_{\delta,\ n}\left|S_{i,j}\right|} \quad (10)$$

Phase 2: Recommendation Generation. For every active user u_a, predicted rating P(u, i) is calculated for each item i ε I not rated by the active user u using the equation (11). Recommendations are generated by computing the sum of the predicted rating related to the items that are similarly weighted by the similarity score and normalized by using the sum of the similarity values.

$$Sim(i,j) = \frac{\sum_{u\epsilon U_{i,j}}^{\alpha}\left(R_{u,i} - \bar{R_\iota}\right)\left(R_{u,j} - \bar{R_J}\right)}{\sqrt{\sum_{u\epsilon U_{i,j}}^{\alpha}\left(R_{u,i} - \bar{R_\iota}\right)^2}\sqrt{\sum_{u\epsilon U}^{\alpha}\left(R_{u,j} - \bar{R_J}\right)^2}} \quad (11)$$

The computational complexity of the TACF approach after clustering is O(KN) + O($\alpha^2\delta$). Here α is neighborhood size and δ is the recent transactions of the user. As the size of α is 20 the complexity is reduced to O(KN) + O(δ). This method is computationally efficient when compared to traditional K-means and IBCF method which is O(KN) + O(n^2m). Here n is the number of items and m is the number of users.

Baseline Estimate

Ranking. In CF patterns or hidden feature can be extracted depending on the user rating behavior. The user may be biased in his rating depending upon their personal preferences and certain items may receive better ratings than the others. All these impacts were clearly captured by Koren (2009) as shown in equation 12 using baseline estimate.

$$u = \mu + b_i + b_u \quad (12)$$

Here u is the unknown rating, μ is the average rating for the dataset, b_i and b_u are the items and user bias calculated as in equation 13 and equation 14.

$$b_i = \frac{\sum_{u \epsilon N}(r_{u,i} - \mu)}{|\{u|(u,i)\ \varepsilon\ N\}|} \quad (13)$$

$$b_i = \frac{\sum_{u \epsilon N}(r_{u,i} - \mu)}{|\{u|(u,i)\ \varepsilon\ N\}|} \quad (14)$$

These parameters are added up to the top-k items and ranked again. These improvements are shown in the experiment section in MAP as TACF + Ranking.

Algorithm 1: Time Adaptive Cluster Method

Input: An active user u_a ε U,
R: The user-item rating matrix, Rank, numIterations,
nClusters, nIterations, nRuns, top-K items.

Output: The items with top-k highest rating $\{i_{r1}, i_{r2}, i_{r3}, \ldots, i_{rk}\}$

1: P ← 0, Q ← random initialize
2: repeat
3: **for** row u ← 1, m **do**
4: $p_u = (Q^T Q + \lambda I)^{-1} Q^T r_u$
5: **end for**
6: **for** column i ← 1, n **do**
7: $q_i = (P^T P + \lambda I)^{-1} P^T r_i$
8: **end for**
9: **until** max iterations
10: $K_p = \sum_{i=1}^{n} \sum_{j=1}^{c} u_{i,j}(p_{i,t})$
11: APPLY LCFDN algorithm on each User Cluster
12: $r' = \mu + b_i + b_u$
13: $r''_{u,i} = u + r'$
14: SORT IN DECREASING ORDER

Algorithm 1 that explains the pseudocode of time-aware scalable neighborhood consists of 3 phases: In the first phase, the rating matrix R, rank and number of iteration that the system that converges are taken as inputs and a matrix factorization model is generated. The matrix R is decomposed into user Factors and item Factors. Each item is represented in some latent factor. The item factor from ALS model is given as input to the time adaptive K-means clustering method. Items are segmented on the basis of the rating behavior of the user by incorporating time.

TACF has better scalability and high performance when compared to traditional methods as item similarity is computed only within the item clusters. The computational complexity is expensive for TACF method as it is a combination of two models. The model is created offline. Since items represent the rating behavior pattern, they produce highly relevant recommendations.

EXPERIMENT ANALYSIS

This section enlists the experiments conducted on the benchmark MovieLens (Harper, & Konstan, 2016) dataset which includes one million ratings from 6040 users on 3900 movies and one lakh ratings from 943 users and 1682 movies. The ratings are represented in numerical form (1~5) along with the rating timestamp. All experiments reported in this section were performed on the machines with Intel(R) Xeon(R)2 CPU 3.36GHz and 32GB RAM and were implemented using Spark. Precision @ k and Mean Square Error (MSE) has been used to evaluate the TACF.

Mean Square Error

The prediction quality of the CF methods is often measured using the statistical accuracy metric MSE (Pentreath, 2015). It is defined as the sum of the squared errors divided by the number of observations as given in equation 15. The lower MSE means high-quality predictions.

$$\mathrm{MSE} = \frac{\sum_{i,j \epsilon X} \left| r_{ij} - r'_{ij} \right|^2}{|X|} \tag{15}$$

where r_{ij} is the actual rating given by user i for an item j, r'_{ij} is the predicted rating

Figure 4 compares the MSE value of ALS, UBCF, IBCF, TSCF, and TACF. MSE values of TACF is found to be much lower compared to other methods. Table 2 represents the ratings predicted by different algorithms for the random user for ml-100k. The predicted values of TACF is more desirable when compared to the other methods. Sparsity helps the K-means produces large deviations from the original ratings. ALS predicts ratings on the basis of the latent factor model. Hence, there are less deviation and lower MSE.

Figure 5 represents the MSE value for the ml-1m dataset. The inference from the figure is that the proposed method TACF excels ALS, UBCF, IBCF and TSCF method. Thus, the proposed method TACF has shown the ability to provide better prediction compared to other traditional models. Similarly, Table 3 represents the predicted value for each movie by random users for ml-1m datasets. There is a greater deviation in K-means and the proposed method TACF shows lower MSE when compared to other methods.

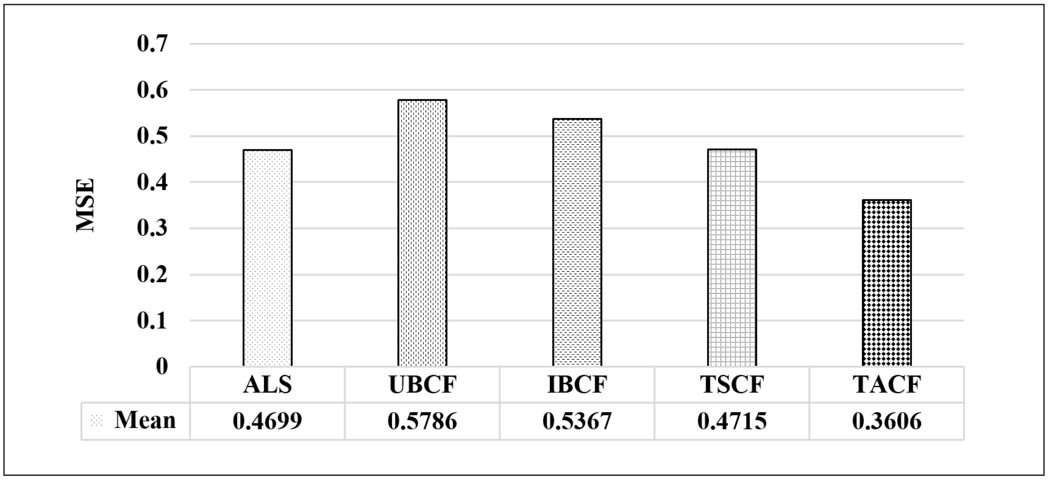


Figure 4. MSE for ml-100k

Table 2
The Predicted value for movies for ml-100k

Random Users	Rating	ALS	UBCF	IBCF	TSCF	TACF
u1	3	4.2455	2.9868	3.2312	3.5517	3.9646
u2	3	3.0646	3.9836	3.3425	3	3
u3	4	4.3988	3.6547	2.6543	3.5813	3.9905
u4	4	4.3993	3.1234	3.4587	3.9651	4.1532
u5	5	4.5889	3.6978	4.3	3.4777	4.7833

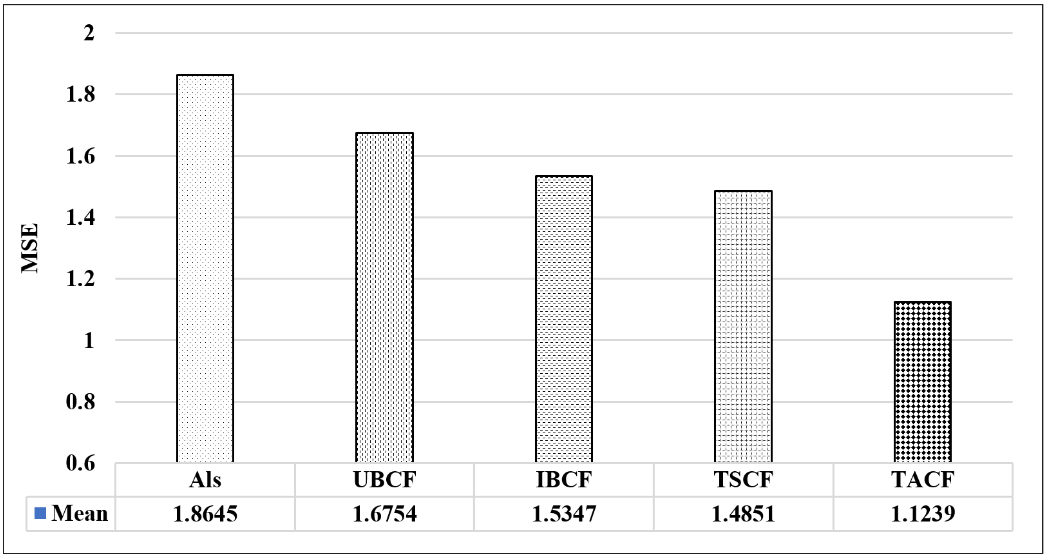


Figure 5. MSE for ml-1m

Table 3
The Predicted value for movies for ml-1M

Random Users	Rating	ALS	UBCF	IBCF	TSCF	TACF
u1	5	3.3576	3.789	3.546	3.6754	3.9687
u2	5	3.2735	4.099	4.2342	4.2034	4.2231
u3	4	3.7959	4.234	3.987	3.7687	3.8687
u4	5	3.746	4.167	4.342	4.407	4.608
u5	5	3.5607	3.774	4.234	4.2314	4.1299

The Mean Average Precision

It is the popular performance measure used for measuring the significance of the top-k items. It is the mean of the average precisions at K (Pentreath, 2015) as indicated in equation 16. Higher MAP values predict a higher quality of the recommendation system.

$$\mathrm{MAP} = \frac{\sum_{q \epsilon Q} \mathrm{AP}(q)}{|Q|} \tag{16}$$

AP is average precision for each query q, |Q| is mean of the queries.

Figure 6 shows the comparison of the MAP values of the ALS, UBCF, IBCF, TSCF, TACF and TACF+Baseline method for ml-100k. TACF method exceeds the UBCF by 65.95%, 69.24%, 89.95%, IBCF by 32.72%, 31.91% and 34.86% and TSCF by 23.73%, 25.63% and 30.24% at top-3, top-5 and top-7 respectively. The percentages calculated and given in Table 4 are meant for a clear illustration.

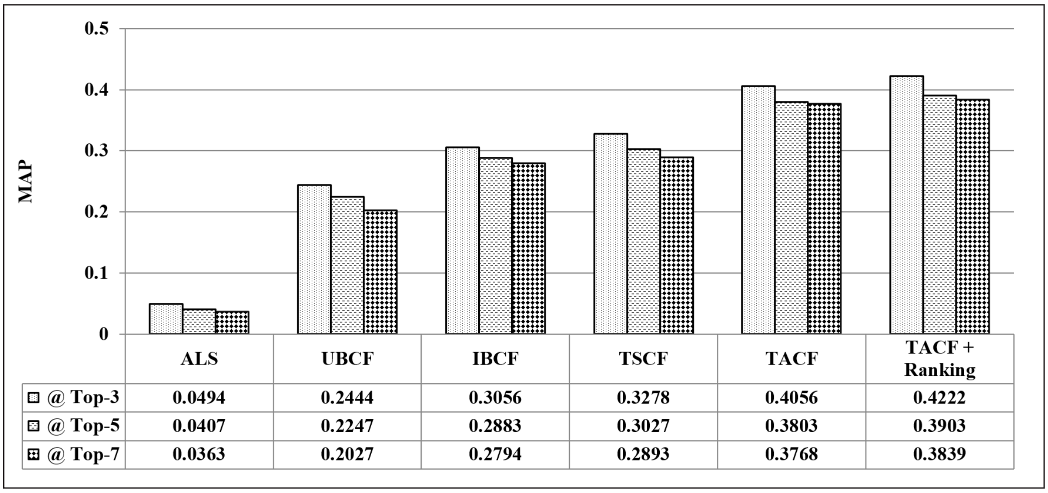


Figure 6. MAP for ml-100k

Table 4
The Percentage Values that excels other methods in MAP for ml-100k

MAP	Top-3	Top-5	Top-7
TACF/ UBCF	65.95 %	69.24 %	85.89 %
TACF/ IBCF	32.72 %	31.91 %	34.86 %
TACF/ TSCF	23.73 %	25.63 %	30.24 %

Figure 7 displays the MAP values for the ml-1M dataset. TACF exceeds UBCF by 84.62%, 50.57% and 37.02%, IBCF by 63.62%, 46.19% and 41.78%, TSCF by 38.44%, 38.53% and 24.36% respectively at top-3, top-5 and top-7. The MAP values of TACF are better compared to other methods. The MAP value decreases with an increase in the number of K.

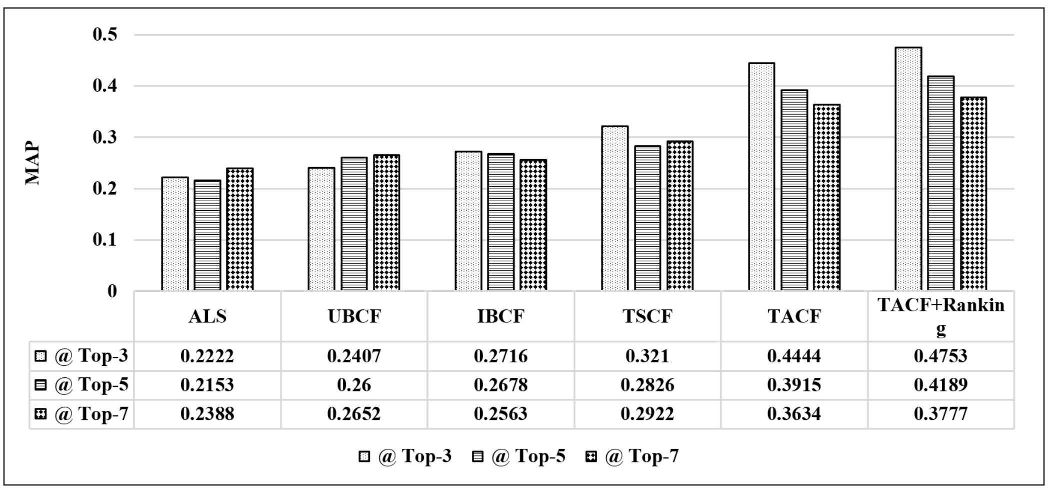


Figure 7. MAP for ml-1m

Table 5 is a clear illustration of MAP values for the ml-1m dataset. The TACF perform better than the other state-of-art methods ALS, UBCF, IBCF and TSCF in MSE and MAP. Thus, the top-k items recommended by the proposed method would satisfy the user's expectation with the recommendation for the items based on the current context of the user.

Table 5
The Percentage Values that excels other methods in MAP for ml-1m

MAP	Top-3	Top-5	Top-7
TACF/ UBCF	84.62 %	50.57 %	37.02 %
TACF/ IBCF	63.62 %	46.19 %	38.53 %
TACF/ TSCF	38.44 %	38.53 %	24.36 %

Execution Time

Figure 8 and Figure 9 show the execution time of different algorithms for ml-100k and ml-1M dataset. The proposed TACF method is computationally faster than traditional methods. The performance of the proposed system will not degrade even if there is a large number of ratings in the rating matrix, as this system considers only 2% of the recent transactions for each user.

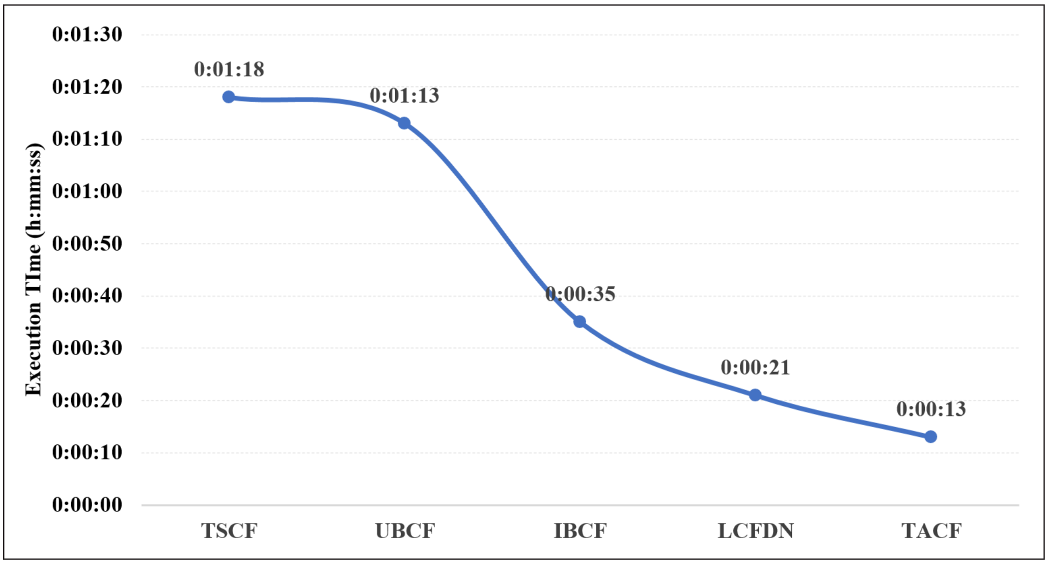


Figure 8. Execution time for generating Recommendation for ml-100k

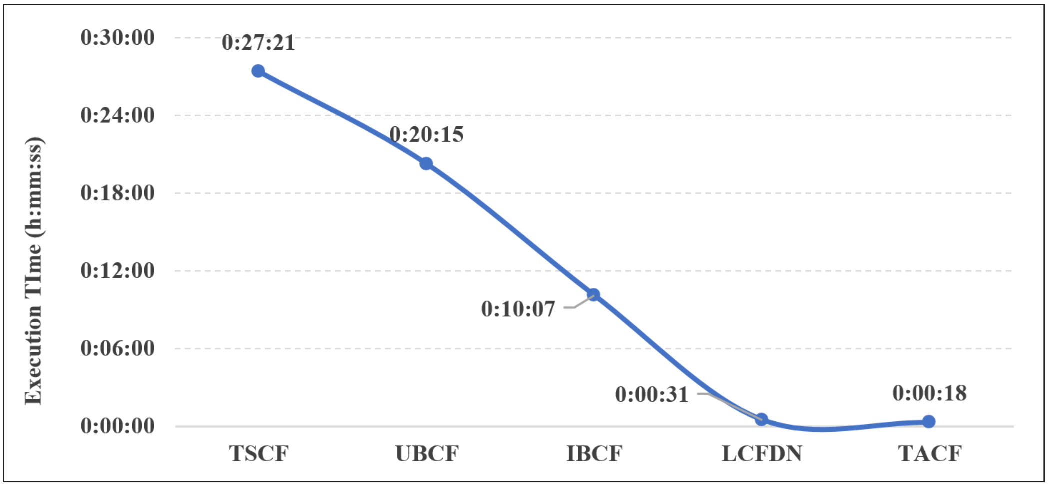


Figure 9. Execution time for generating Recommendation for ml-1M

CONCLUSION

This paper presents a time-adaptive recommendation approach on item subgroups that exploits the recency of the user's transaction to address the data sparsity problem. The items are clustered based on the item latent factors and time is used as additional information in generating recommendation on each subgroup. Experimental results show that using time in item subgroups is a promising way to improve the top-k recommendation performance when compared with the other state-of-the-art collaborative filtering methods. Future work can be extended by forming better user or item subgroups by adding additional information based on context.

ACKNOWLEDGEMENT

We are immensely grateful to the anonymous reviewers for their valuable comments and suggestions to improve the quality of the article.

REFERENCES

Adomavicius, G., & Tuzhilin, A. (2005). Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering, 6*(17), 734-749.

Ba, Q., Li, X., & Bai, Z. (2013, May 23-25). Clustering collaborative filtering recommendation system based on SVD algorithm. In *2013 4th IEEE International Conference on Software Engineering and Service Science (ICSESS)* (pp. 963-967). Beijing, China.

Bell, R. M., & Koren, Y. (2007). Lessons from the Netflix prize challenge. *Acm Sigkdd Explorations Newsletter*, *9*(2), 75-79.

Billsus, D., & Pazzani, M. J. (1998, July 24-27). Learning Collaborative Information Filters. In *Proceedings of the 15th International Conference on Machine Learning, Madison, WI,* (Vol. 98, pp. 46-54). Morgan Kaufmann, San Francisco.

Chen, J., Fang, J., Liu, W., Tang, T., Chen, X., & Yang, C. (2017, May 29-June 2). Efficient and portable ALS matrix factorization for recommender systems. In *IEEE International conference on Parallel and Distributed Processing Symposium Workshops (IPDPSW)* (pp. 409-418). ake Buena Vista, FL, USA.

DeCoste, D. (2006, June 25-29). Collaborative prediction using ensembles of maximum margin matrix factorizations. In *Proceedings of the 23rd international conference on Machine learning* (pp. 249-256). Pittsburgh, Pennsylvania, USA.

Desrosiers, C., & Karypis, G. (2011). A comprehensive survey of neighborhood-based recommendation methods. In F. Ricci, L. Rokach, B. Shapira & P. Kantor (Eds.), *Recommender systems handbook* (pp. 107-144). Boston, Massachusetts: Springer.

Ding, Y., & Li, X. (2005, October 31-November 5). Time weight collaborative filtering. In *Proceedings of the 14th ACM international conference on information and knowledge management* (pp. 485-492). Bremen, Germany.

Guo, G., Zhang, J., & Yorke-Smith, N. (2015). Leveraging multiviews of trust and similarity to enhance clustering-based recommender systems. *Knowledge-Based Systems*, *74*, 14-27.

Harper, F. M., & Konstan, J. A. (2016). The movielens datasets: History and context. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, *5*(4), 1-19.

Heckerman, D., Chickering, D. M., Meek, C., Rounthwaite, R., & Kadie, C. (2000). Dependency networks for inference, collaborative filtering, and data visualization. *Journal of Machine Learning Research*, *1*(Oct), 49-75.

Herlocker, J., Konstan, J. A., & Riedl, J. (2002). An empirical analysis of design choices in neighborhood-based collaborative filtering algorithms. *Information Retrieval*, *5*(4), 287-310.

Hu, Y., Peng, Q., Hu, X., & Yang, R. (2015). Time-aware and data sparsity tolerant web service recommendation based on improved collaborative filtering. *IEEE Transactions on Services Computing, 8*(5), 782-794.

Jiang, S., Qian, X., Mei, T., & Fu, Y. (2016). Personalized travel sequence recommendation on multi-source big social media. *IEEE Transactions on Big Data*, *2*(1), 43-56.

Koren, Y. (2009, June 28-July 01). Collaborative filtering with temporal dynamics. In *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 447-456). Paris, France.

Li, Q., & Kim, B. M. (2003, October 13-17). Clustering approach for hybrid recommender system. In *Proceedings IEEE/WIC International Conference on Web Intelligence (WI 2003)* (pp. 33-38). Halifax, NS, Canada.

Luo, X., Liu, H., Gou, G., Xia, Y., & Zhu, Q. (2012). A parallel matrix factorization based recommender by alternating stochastic gradient decent. *Engineering Applications of Artificial Intelligence*, *25*(7), 1403-1412.

Ma, X., Lu, H., Gan, Z., & Zhao, Q. (2016). An exploration of improving prediction accuracy by constructing a multi-type clustering based recommendation framework. *Neurocomputing*, *191*, 388-397.

Meng, S., Dou, W., Zhang, X., & Chen, J. (2014). KASR: a keyword-aware service recommendation method on MapReduce for big data applications. *IEEE Transactions on Parallel and Distributed Systems*, *25*(12), 3221-3231.

Miyahara, K., & Pazzani, M. J. (2000, August). Collaborative filtering with the simple Bayesian classifier. In R. Mizoguchi & J. Slaney (Eds.), *PRICAI 2000 Topics in Artificial Intelligence-PRICAI 2000. Lecture Notes in Computer Science* (pp. 679-689). Heidelberg, Germany: Springer.

O'Connor, M., & Herlocker, J. (1999, August). Clustering items for collaborative filtering. In *Proceedings of the ACM SIGIR workshop on recommender systems* (Vol. 128). Berkeley, California.

Panniello, U., Tuzhilin, A., & Gorgoglione, M. (2014). Comparing context-aware recommender systems in terms of accuracy and diversity. *User Modeling and User-Adapted Interaction*, *24*(1-2), 35-65.

Park, H., Jung, J., & Kang, U. (2017, December 11-14). A comparative study of matrix factorization and random walk with restart in recommender systems. In *Big Data (Big Data), 2017 IEEE International Conference on* (pp. 756-765). Boston, Massachusetts.

Pentreath, N. (2015). Building a Recommendation Engine with Spark. In N. Pentreath (Ed.), *Machine Learning with Spark* (pp. 83-116). Birmingham UK: Packt Publishing Ltd.

Ren, Y., Li, G., Zhang, J., & Zhou, W. (2013). Lazy collaborative filtering for data sets with missing values. *IEEE transactions on cybernetics*, *43*(6), 1822-1834.

Sarwar, B. M., Karypis, G., Konstan, J., & Riedl, J. (2002a, December 27-28). Recommender systems for large-scale e-commerce: Scalable neighborhood formation using clustering. In *Proceedings of the Fifth International Conference on Computer and Information Technology* (Vol. 1, pp. 291-324). Dhaka, Bangladesh.

Sarwar, B., Karypis, G., Konstan, J., & Riedl, J. (2002b, December 27-28). Incremental singular value decomposition algorithms for highly scalable recommender systems. In *Fifth International Conference on Computer and Information Science* (pp. 27-28). Dhaka, Bangladesh.

Sarwar, B., Karypis, G., Konstan, J., & Riedl, J. (2001). Item-based collaborative filtering recommendation algorithms. In *Proceedings of the 10th international conference on World Wide Web* (pp. 285-295). New York, NY, USA: Association for Computing Machinery.

Sarwat, M., Levandoski, J. J., Eldawy, A., & Mokbel, M. F. (2014). LARS*: An efficient and scalable location-aware recommender system. *IEEE Transactions on Knowledge and Data Engineering*, *26*(6), 1384-1399.

Su, X., & Khoshgoftaar, T. M. (2006, November 13-15). Collaborative filtering for multi-class data using belief nets algorithms. In *18th IEEE International Conference on Tools with Artificial Intelligence, 2006 (ICTAI'06)* (pp. 497-504). Arlington, VA, USA.

Suganeshwari, G., & Ibrahim, S. P. S. (2016). A survey on collaborative filtering based recommendation system. In V. Vijayakumar & V. Neelanarayanan (Eds.), *Proceedings of the 3rd International Symposium on Big Data and Cloud Computing Challenges (ISBCC–16')* (pp. 503-518). Cham, Switzerland: Springer.

Suganeshwari, G., Ibrahim, S. P. S., & Li, G. (2018). Lazy collaborative filtering with dynamic neighborhoods. *Information discovery and delivery*, *46*(2), 95-109.

Sun, L., Michael, E. I., Wang, S., & Li, Y. (2016, December 15-18). A Time-Sensitive Collaborative Filtering Model in Recommendation Systems. In *IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)* (pp. 340-344). Chengdu, China.

Takács, G., Pilászy, I., Németh, B., & Tikk, D. (2008, December 15-19). Investigation of various matrix factorization methods for large recommender systems. In *IEEE International Conference on Data Mining Workshops, 2008 (ICDMW'08)* (pp. 553-562). Pisa, Italy.

Tan, S., Bu, J., Chen, C., Xu, B., Wang, C., & He, X. (2011). Using rich social media information for music recommendation via hypergraph model. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, *7*(1), 1-22.

West, J. D., Wesley-Smith, I., & Bergstrom, C. T. (2016). A recommendation system based on hierarchical clustering of an article-level citation network. *IEEE Transactions on Big Data*, *2*(2), 113-123.

Xue, G. R., Lin, C., Yang, Q., Xi, W., Zeng, H. J., Yu, Y., & Chen, Z. (2005, August 15-19). Scalable collaborative filtering using cluster-based smoothing. *In Proceedings of the 28th Annual International ACM SIGIR conference on Research and Development in Information Retrieval* (pp. 114-121). New York, NY, USA.

Yu, H. F., Hsieh, C. J., Si, S., & Dhillon, I. S. (2014). Parallel matrix factorization for recommender systems. *Knowledge and Information Systems*, *41*(3), 793-819.

Pertanika J. Sci. & Technol. 27 (4): 1803 - 1820 (2019)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Criminality Recognition using Machine Learning on Malay Language Tweets

Nurul Hashimah Ahamed Hassain Malim*, Saravanan Sagadevan and Nurul Izzati Ridzuwan

School of Computer Sciences, Universiti Sains Malaysia, 11800 USM, Pulau Pinang, Malaysia

ABSTRACT

A large scale of investigation had been carried out to predict the personality, or in precise, the behaviour of online users through user-generated texts, such as Tweets and status messages. Nevertheless, only a handful of machine learning (ML) studies have applied the personality model to assess criminality behaviour, particularly within the context of Malay social network messages. Based on the concept of sentiment valence, this study annotated a list of Malay Tweets that might be subjected to crime or illicit messages from the stance of Psychoticism trait. Consequently, the supervised-based text classification method was conducted by using Naïve Bayes (NB), Sequential Minimal Optimisation (SMO), and Decision Tree (DT) on Tweets using several features determined via Chi Square (x^2). The analyses outcomes signified that SMO outperformed other classifiers insignificantly by achieving 92.85% of accuracy. Based on x^2, several swear terms, such as *bontot*, *melancap*, and *kote*, displayed significant correlation with Psychoticism Tweets due to the nature of the trait that has been subjected to criminality behaviour, for instance, aggressive and antisocial attributes. The findings illustrate the possibilities to adapt several personality aspects in order to enhance the effectiveness in detecting illicit social network messages.

Keywords: Machine learning, Malay tweets, personality detection, text mining

ARTICLE INFO

Article history:
Received: 07 February 2019
Accepted: 04 June 2019
Published: 21 October 2019

E-mail addresses:
nurulhashimah@usm.my (Nurul Hashimah Ahamed Hassain Malim)
saravanan_18@student.usm.my (Saravanan Sagadevan)
izzati.ucom13@student.usm.my (Nurul Izzati Ridzuwan)
* Corresponding author

INTRODUCTION

In earlier time, Sir Francis Galton had hypothesised that natural language terms might represent personality differences in humankind, while Allport et al. (1936) claimed that almost 18,000 English terms could represent personality, and Hofstee (1990) suggested that nouns, sentences,

ISSN: 0128-7680
e-ISSN: 2231-8526

and actions might mirror one's personality (Gerald et al., 2003). At present, the vast participation of online users in social network activities has contributed to unprecedented textual resources, such as tweets that are rich in personal discourses. For example, Twitter has become a source that generates a substantial amount of user data, wherein daily active users of the platform reached 157 million as of second quarter of 2017 and approximately 500 million tweets were shared each day among users (Kursuncu et al., 2018). The massive number of user activities on social networks offers valuable insights about one's behaviours, experiences, opinions, and interests. The insights are fundamentally directed via psychology, or in specific, personality that advocates both emotional aspects and characteristics of a human being.

As the domain of data science is expanding every passing day, interest towards recognising and understanding personality traits of virtual users via user-generated content, especially from text data, has grown rapidly in recent years. Since the study by Argamon et al., (2005) that used functional lexical features from student essays to automatically predict Extraversion and Neuroticism traits, many empirical investigations have been conducted to reckon human traits from digital data. Meanwhile, Oberlander and Nowson (2006) opened a new chapter in evaluating the personality amidst social media users through lexical analysis. The investigation revealed that the combination of attributes that derived from language models yielded good outcomes in classifying the personality of weblog users. Subsequently, many scholars have begun exploring the latent representation of personality characteristics embedded in texts.

One of the most prominent efforts to evaluate the personality of online users have been initiated by using *myPersonality* dataset and a common benchmark in a Workshop on Computational Personality Recognition (Shared Task) (Celli et al., 2013). The workshop refers to gold standard that evaluates the performance of Machine Learning (ML) algorithms using Big Five Personality traits, namely Openness to Experience, Conscientiousness, Extroversion, Neuroticism, and Agreeableness, to represent the nature of human being characteristics. The empirical review displayed that many personality recognition studies have embraced the description defined in Big Five model. For instance, based on the Big Five model, Peng et al., (2015) employed the Chi Square and Recursive Feature Elimination (RFE) selection methods to detect personality from Chinese texts, while Aalderks (2014) performed Latent Semantic Analysis to determine significant features so as to identify non-cognitive personality in post-secondary student essays. Shally (2014) substantially investigated the online behaviours of Facebook and LinkedIn users using Big Five traits, while Correa et al., (2010) and Guadagno et al., (2008) reported that users who scored highly on Neuroticism frequently used more social network platforms, such as Facebook and Twitter.

Additionally, several studies (Gerald et al., 2003; Oberlander & Nowson, 2006; Peng et al., 2015) have presented evidence of strong inter-correlation across personality-behaviour-linguistics areas. The strong inter-correlation among these areas should be further explored by incorporating relevant domain knowledge into other relevant aspects, such as linguistics-based criminality recognition. Criminality writing recognition within the context of personality may significantly assist the detection of language-based cybercrime activities, such as cyber bully and cyber harassment (Farshad et al., 2016). Nevertheless, only a handful of studies have adapted other personality models or have explored the dynamics of text messages from the perspective of crime.

Although no investigation has explored the correlation between crime and texts in the context of personality, several studies (Sagadevan et al., 2015; Saravanan, 2016) have assessed the representation of Psychoticism trait from the Psychoticism, Extraversion, and Neuroticism (PEN) Model, which is typically aggregated to criminal behaviours (Kamaluddin et al., 2015). As Psychoticism trait represents certain behaviours, such as aggressiveness and interpersonal hostility that are naturally linked with negativity and crimes (Kamaluddin et al., 2015), the integration of PEN framework with other pipeline concepts, such as sentiment valences and Automatic Personality Perception (APP), sheds light on the nature of criminality digital writings. The description of PEN Model traits and their specific characteristics are presented in Table 1.

Table 1
PEN dimensions (Allport, 1961)

Trait	Characteristics
Extraversion	Sociable, lively, active, assertive, sensation seeking, carefree, dominant, surgent and venturesome
Neuroticism	Anxious, depressed, guilt feelings, low self-esteem, tense, irrational, shy, moody and emotional
Psychoticism	Aggressive, cold, egocentric, impersonal, impulsive, antisocial, unempathetic, creative and tough-minded

Past studies have probed into the availability of open source English text messages from Facebook (Celli et al., 2013) and Twitter (Go et al., 2009) to assess proximal cues, which aggregated to a well-known criminal trait called Psychoticism by cross-validating the personality description provided in the psychological self-assessment report (derived from Facebook). Initially, Part-Of-Speech (POS) Tagging was implemented on cleaned datasets to identify adjectives, nouns, and verbs that brought sentiment polarity, wherein those modalities served as subject of study in the public questionnaire (Sagadevan et al., 2015) to gather general perception regarding sentiment valences and their association with PEN Model traits. The execution of questionnaire to identify the effects of words on personality trait is a perception-based approach that falls under APP, which can determine the co-relationships between sentiment valences of a term and personality traits.

The APP refers to an Automatic Personality Detection approach that emphasises on predicting personality attributes based on observable behaviours, such as types of words used in writings. Albeit the prediction is based on other observations and perceptions of the public in influencing the personality in social interactions (Mohammadi & Vinciarelli, 2012), it is the predisposition of humans to comprehend the behaviours of others based on the observation of their behaviours in everyday life (Agarwal, 2014).

Valence refers to one of the Dimensional Views proposed by Osgood et al., (1957), which can be applied to measure the level of pleasantness and unpleasantness of each word (Bradley & Lang, 1999). Subsequently, the pre-processed dataset annotated using frequency (Schwartz, 2013) of sentiment words was categorised based on valences schema (Table 2) and applied feature selection prior to transformation into vectors that represented stream of strings. Feature selection is a method that reduces data dimensionality by selecting a subset of attributes from a bigger pool of inputs to devise a prediction system that can preserve the original information as much as possible.

Based on supervised learning, many ML classification experiments have been conducted on multiple types of language models using several prominent algorithms, such as Sequential Minimal Optimisation (SMO), Naïve Bayes (NB), K-Nearest Neighbour (KNN), and Decision Tree (DT). Saravanan (2016) reported that SMO and Unigram language model yielded promising outcomes, when compared to other classifiers, and forwarded suggestions that quadratic-based optimisation classifiers could perform well on high dimensional inputs and single-text attributes might channel valuable information towards automatic learning process.

To the best of the authors' knowledge, no study has used the Malay social media text messages to infer personality traits of users, especially within the context of criminality behaviours. Most of the past empirical studies that used Malay texts placed more focus on the typical linguistics domains, such as sentiment analysis (Chekima & Alfred, 2018; Al-Saffar et al., 2018; Al-Moslmi et al., 2017; Darwich et al., 2016) and machine translation (Wang et al., 2015).

Since the earlier study conducted by Saravanan (2016), which focused on English social network messages, yielded promising findings, the efforts of this present study are extended to study the representation of Malay Tweets in the context of Psychoticism and to measure the performance of ML algorithms in predicting traits-based instances. In this study, the authors assumed that perceptions towards semantic of lexical, which had been typically linked with sentiment polarity, might serve as an indicator to classify the sentences based on PEN Model traits. The following sections elaborate the methods, the outcomes, and the discussions of this study.

MATERIALS AND METHODS

The methodology in this study is composed of data collection and several pre-processing techniques, such as data cleaning, annotations, and eventually, automatic classification by three ML algorithms (NB, SMO & DT). Figure 1 illustrates the methodology of this study.

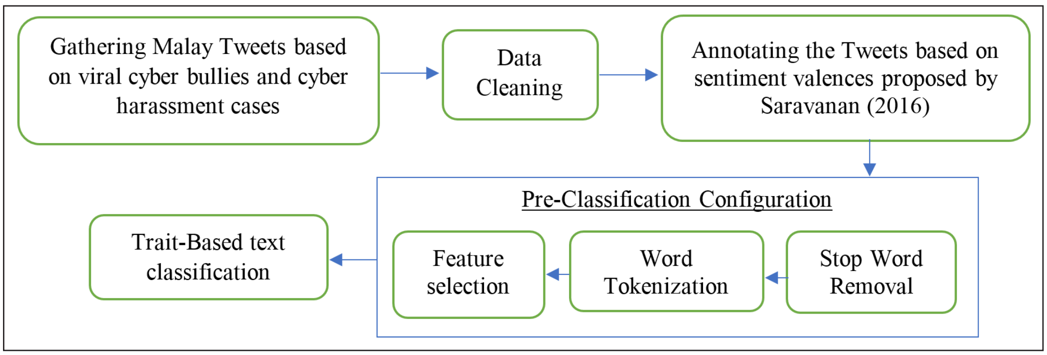


Figure 1. Methodology

Prior to data collection, the trending topics on Twitter that turned viral due to harassment or bully of other Twitter members were observed. Next, the types of words used by the account owners were randomly observed and studied. Since the lexicon resources for Malay language is limited and most of them have been readily available off-the-shelf classifiers designed for English texts (Darwich et al., 2016), the seed words applied in a previous study were translated to the Malay language and served as hints to identify the possible true cyber bullies and cyber harassment. The seed words and their sentiment valences are presented in Figure 2. Approximately one month was taken by the researchers to assess the

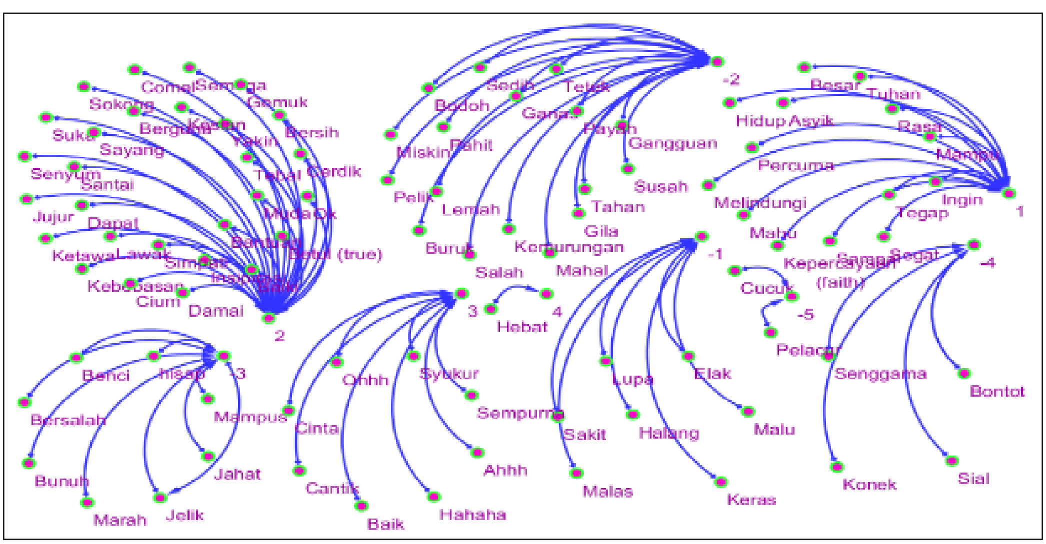


Figure 2. Sentiment valences-based seed words

messages posted in the selected accounts before analysing their tweets using *Tweepy* and labelling them in accordance to the schema proposed by Saravanan (2016). As for the non-criminal message corpuses, the researchers randomly crawled tweets from other trending topics, such as fashion, business, and news, apart from inspecting the context of the data in order to ensure that they were excluded from harassment and bully relationship. Due to time constrain, the data crawling merely yielded small scales of corpus that comprised 82,293 words, in which each instance consisted of about 2,939 words in average.

Due to the nature of the social networking tweets that is in the form of unstructured and unformatted (Salloum et al., 2017), data cleaning process was performed to normalise the structure of the tweets to meet standard forms. As the amount of the collected data had been small, the data cleaning process was manually conducted and particularly discarded meaningless strings, such as Uniform Resource Locator (URL), punctuation marks, symbols, numbers, and word standardisation from malleable to regulated forms (mls to malas). As a result, the data were annotated using the schema proposed by Saravanan (2016). Based on the schema, first, the words in the dataset were labelled as either positive or negative, and next, the words were matched with their sentiment by adhering to the rules presented in Table 2.

Table 2
Sentiment/valence categorisation

Trait	Sentiment	Valence
Extraversion	Positive	1 to 5
Neuroticism	Lower Negative	(-1) to (-3)
Psychoticism	Higher Negative	(4) to (-5)

Based on the PEN model, the Extraversion trait is significantly aligned to positive behaviours, whereas Neuroticism and Psychoticism lean towards negative characteristics (Kamaluddin et al., 2015). By adhering to this concept, all the sentiment terms were ranked based on frequency and were categorised according to the PEN model traits in identifying the sentiment strength of the collected tweets. As the term frequency method is a prominent technique applied in text mining to determine sentiment intensity of textual streams (Schwartz et al., 2013), this present study embraced the statistical approach to calculate both the intensity and the tendency of the users towards the relevant PEN model traits. As the goal of this study is to evaluate the representation of Psychoticism trait in Malay tweets, the term frequency ranking process was initiated from higher negative to positive as the nature of the online texts is more positive and negative, as well as the fact that the latter carry more information than the former (Garcia et al., 2012). Subsequently, the average word usage of each trait category was determined by using the formula given in (1). The average word counting method (1) is a common technique employed to bridge the connection between language and psychological variables that belongs to manually-constructed classes of language (Schwartz, 2013).

$$Average\ word\ usage = \frac{Number\ of\ trait\ related\ words}{Total\ number\ of\ users\ in\ each\ traits} \quad (1)$$

Although an earlier study have reported that as minimum as five higher negative words usage in average matched the cross-check correlation between Psychoticism and Agreeableness + Neuroticism (Van Dam et al., 2005), this study revealed that the average usage of higher negative words was seven. Nonetheless, the instances were only labelled as Psychoticism if the document consisted as minimum as 14 higher negative words, mainly due to the small volume of dataset. The stipulated figure served as a guide to represent the data as structurally compact and to offer adequate volume of Psychoticism instances so as to hinder overfitting problem during ML classification. The annotation for Neuroticism instances adhered to formula (1), while the average Neuroticism words usage was set to 48. As the PEN Model merely consists of three global traits, the other instances were assumed to be more adjacent to positive polarity and were annotated as Extraversion. Eventually, the Psychoticism, Neuroticism and Extraversion classes contained 8696, 36,573, 37,024 number of words respectively. The examples of the Tweets representing each of the traits illustrated in Table 3.

Table 3
Example of tweets each classes of trait

Class	Example of Tweets
Psychoticism	*Thanks* gemok pun dah tunang Hang Korek puki pakai jari ja ka tiap malam
Neuroticism	Susah jaga orang tua hampa tau ka
Extraversion	*Dear students* Belajar ibarat naik bukit Bila kau dah sampai puncak kau akan nampak permandangan yg sgt cantik Jangan mudah putus asa

The instances from the three traits were represented as a set of document *d,* wherein each d_i signified a sequence of sentence *s* and was classified based multi-classes classification, primarily because the PEN model is composed of three global traits. The training document *d* was transformed into vectors by using the Bag of Words (BOW) technique. In precise, the string to vector transformation involved removal of stop words, such as *ini, itu, sana, dari*, and *yang*, while string tokenisation was conducted on unigram attributes because the scale of the present dataset was small, along with the high possibilities of individual variances to articulate using single terms (Mairesse & Walker, 2011). After that, in the attempt of minimising the dimensionality and to enhance the detection performance, several relevant significant attributes were selected based on Information Gain (IG). The IG refers to a function of probability distribution initiated from communication theory based on entropy mechanism that measures the uncertainty of random attributes, such as words in textual stream (Mitchell, 1997). The entropy measures the expected reduction caused by partitioning the samples based on the attributes. Simply put, IG calculates the number of bits of information gathered to predict classes by estimating the presence or absence of certain *words* in a document (Gao et al., 2014). In text mining, Entropy measures the information or knowledge encapsulated in vectors mapped from text contents. The lack of information or high entropy value will lead to poor prediction whereas low entropy will assists better prediction process. The IG for a *word (w)* is defined based on (2). Next, automatic classification was performed by using the algorithms stated in the following section based on 10-fold cross-validation resampling procedure due to the limited number of training instances (Refaeilzadeh et al, 2009).

$$IG:(D,w) = Entropy\ (D) - Entropy\ (D|w)\,,$$

$$Entropy\ (D) = \sum_{c \in C} -p(w)\ log_2\ p\ (w) \qquad (2)$$

where, *D* refers to the training document where *w* is the set of all possible attributes (words).

P(w) is a word that is present in the instances that belongs to class $c \in C$.

Machine Learning Algorithms

This study measured the performances of three selected algorithms, namely NB, SMO, and DT, against the majority baseline. The three off-the-shelf algorithms were chosen in this study, since many text mining studies had employed these classifiers in their analyses (Nasa & Suman, 2012; Sujatha & Ezhilmaran, 2013; Kapur et al., 2017). These off-the-shelf algorithms are suitable for automatic classification with small volume of data (Ruparel et al., 2013). The following sub-sections describe each selected machine classifier.

Baseline

After taking example from a prior study (Celli et al., 2013), this present study employed a classifier called Zero Rules (ZeroR) as a majority baseline to measure the performances of each ML classifier applied in this study. The ZeroR predicts the mean (for numeric-type target attribute) or the mode (for nominal-type attribute) of the most commonly found attributes in the datasets and does not apply any rule that works on non-target attributes (Barber, 2012).

Naïve Bayes (NB)

The NB is a very popular probabilistic classifier based on Bayes rules with strong independent assumptions. In other word, a descriptive independent feature model» based on probability will make NB to assume that the presence or absence of a peculiar feature of a class is not related to the presence or absence of other features (Witten et al., 2011). The formula of NB illustrated in (3). Based on the model, the *P(c)* is a prior probability that initially set to *c* whereas *P(d)* indicated the initial probability of the training document *d*. *P(d|c)* referred to the appearance probability of document *d* when *c* is established. The *P(c|d)* is a calculation of posterior probability that represents the confidence of presence of *c* on *d.* As nature of text documents contains multiple attributed transformed through BOW that made the independent assumption on word orders, the *d* represented as word features x_1, $x_2 \ldots x_n$ and denoted as conditional independence $P(x_1,\ldots x_n|c) = P(x_1|c) \cdot P(x_2|c) \cdot \ldots . P(x_n|c)$. Eventually, NB made the decisions of the classes belongs to particular instances based on fractions of times words x_i appears among other *x* in *d* by estimating the maximum likelihood of *c* (Witten et al., 2011).

$$P(c|d) = \frac{P(d|c)P(c)}{P(d)} \tag{3}$$

*Where *c* referred to training classes and *d* is training document.

Sequential Minimal Optimisation (SMO)

SMO is a variant of SVM that optimize a problem iteratively by splitting the problem into a series of sub-problems (Huang & Yan, 2014). The selection of smallest optimization problem involved two Lagrange Multipliers where both must follow the equality constraints and jointly optimized to find the optimal values as well as update the vector machine to reflect the new optimal values. This process was repeated until the convergence criteria was met. To speed up the convergence, SMO use heuristics to select both Lagrange Multipliers to jointly optimize the problem. The problem has been solved when all the Lagrange Multipliers satisfy the Karush–Kuhn–Tucker (KKT) conditions (Huang & Yan, 2014). The working procedure of SMO illustrated in Figure 3. In the first place, the selection of dual Lagrange Multipliers m_1 and m_2 were based on heuristic methods. The first heuristic provides the outer loop to iterate over the entire training set to determine the violations of KKT conditions by each instance. The determination evaluated based on computations of m_2's upper bound, *u,* and lower bound. *l.* If any of the instances violate the KKT condition, then it is eligible for optimization. The second heuristics intended to support another multiplier to maximize the size of the step taken during joint optimization. Consequently, the m_2 need to be update based on Δm_2. The updating process fails if the Δm_2 is smaller than threshold, otherwise, m_1 should be update along with all of the objective functions, f_i, values.

Further, computation conducted to determine the deviation between the output of function and classification target, t. SMO will terminate if the value of t below than threshold. The pair-wise classification mechanism was applied in this study as the training data consisted of instances from three classes (multi-class classification). In order to prevent inter-class generalisation problem, Polynomial kernel functions were applied due to the popularity of the technique in NLP (Goldberg & Elhadad, 2008).

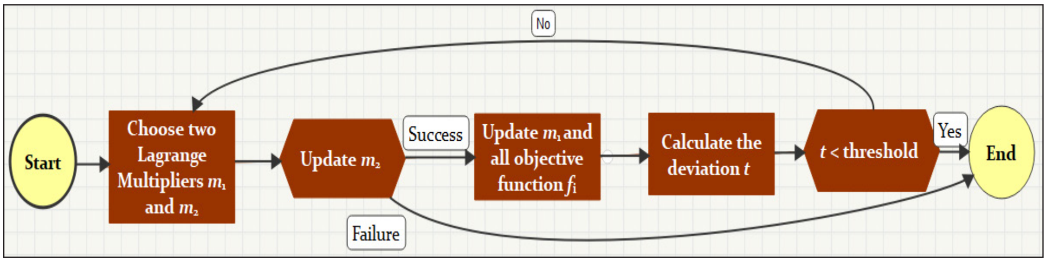


Figure 3. SMO training process

Decision Tree (DT)

Theoretically, DT refers to a tree-based classifier, wherein each branch of nodes denotes a selection between a list of alternatives and each leaf depicting the decisions (Dunham, 2006). In precise, the foundation of the classifier is to determine the structure of the vectored-attributes behaviour for several instances in the classes and newly-generated instances (Korting, 2006). In particular, one of the DT classifiers called J48 was employed to predict the classes in this study. The J48 adhered to the implementation of C4.5 classifier and incorporated several additional functions, such as tackling missing values, DT pruning, and rules derivation (Kaur & Chhabra, 2014). Generally, DT algorithm follows the following steps to classify the instances.

(a) First, J48 develops a DT based on the attribute values of the available training data or from the 10-fold cross-validation.
(b) Second, J48 identifies the attributes that discriminate the various instances with the highest information gain.
(c) Third, the non-ambiguous attributes branches terminate and assign the attributes to the target value.

Evaluation Metric

The accuracy evaluation metrics were applied to measure the performance of ML algorithms in this study. Accuracy is a widely applied evaluation metrics in ML classification (Hossin & Sulaiman, 2015). Typically, ML studies use statistical metrics called confusion matrix to measure the strength of an algorithm in solving the given problem automatically. The confusion matrix comprised four elements of conditions, namely True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN). The TP refers to the proportion of positive cases that are correctly identified, FP is the proportion of negative cases that are incorrectly classified as positive, while TN is the proportion of negative cases that are classified correctly, and FN is the proportion of positive cases that are incorrectly classified as negative (Hossin & Sulaiman, 2015). The accuracy of evaluation metrics is determined based on the mathematical operation on the elements of confusion matrix, as illustrated in formula (4).

$$\frac{TP+TN}{TP+TN+FP+FN} \qquad (4)$$

RESULTS AND DISCUSSION

Based on the methodology described, this study investigated the performance of ML algorithms using supervised learning mechanism and assessed the existing correlation between Neuroticism and Psychoticism, apart from extracting several statistically significant terms linked with Psychoticism trait. The following subsections present the findings of each investigation.

Machine Learning Evaluation

Three ML classifiers, namely SMO, NB and DT, had been employed in this study to evaluate the prediction accuracy on the supervised-based constructed inputs that represented the PEN model traits instances. Based on the outcomes displayed in Figure 4, all the classifiers performed better than the baseline did, with SMO leading the prediction and followed by NB and DT. The outperformance exhibited by SMO is in agreement with the findings reported in previous studies concerning personality recognition (Verhoeven et al., 2013; Saravanan, 2016), which exploited the representation of English texts. Basically, SMO learnt the representation via analytic quadratic programming that decomposed the overall inputs into sub-problems and iteratively optimised the smaller units using Lagrange multipliers until the problem was solved. The goal of SMO is to forward the Lagrange multipliers or alphas that satisfy the actual inherent learning process by identifying the support vectors (Platt, 1998), whereas the role of kernel function is to transform the inputs into a higher dimension and to optimise the sub-problems until convergence prevents large matrix computation (Urmaliya & Singhai, 2013). Hindering computation on large matrix enables SMO to exploit massive volume of sparse text dataset that contains a huge number of zero vector elements.

Meanwhile, NB significantly performed well in this present study, when compared to that reported by Saravanan (2016) that predicted below the baseline upon use of larger instances as the input. The significant prediction displayed by NB is attributable to the small volume of the instances applied in this study, which adjusted the weights for decision boundary, minimised the bias effects, and normalised the magnitude of the weights for strong and weak word dependency classes (Rennie et al., 2003). The exceptional performance exerted by NB contradicts that stated by Raschka (2014), who denoted that the probability-based classifier tended to perform poorly on non-linear classification problems. The significant performance signifies that NB can perform well on non-linear problems if the volume of the training instances can normalise the weights and when the magnitude of the weights is assigned to the decision boundary and the attributes of the classes.

The DT classifier merely achieved 70.96% of accuracy. As the nature of DT is greedy, each separation in the tree was determined based on isolation without considering the possible biases in future tree that may poorly capture the underlying characteristics of training instances. The poor pruning mechanism due to small volume of dataset may cause weak generalisation and potentially directed to underperformance in predicting future points (Bertsimas & Dunn, 2017).

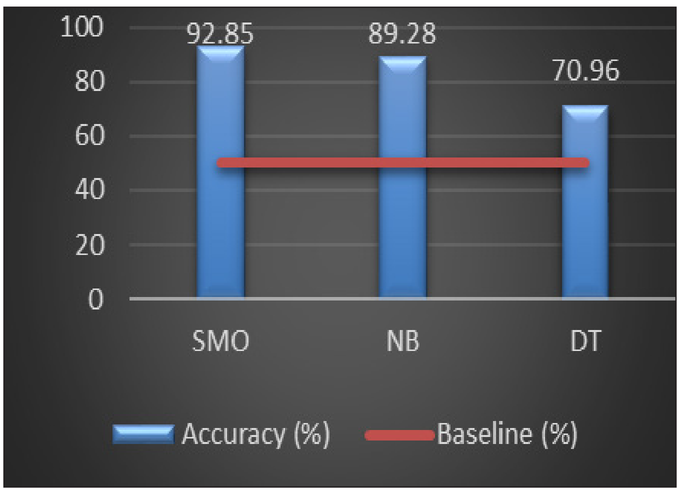


Figure 4. Machine learning classification result

Projection Plot

The analysis was extended to determine the correlations present between Psychoticism, Neuroticism, and Extraversion instances. Projecting the instances in the form of plot graph illustrated an interesting pattern, whereby strong correlation was discovered between the language structures of Psychoticism and Neuroticism (Figure 5). To recap, although both types of respective traits characterised negative behaviours, it was detached based on higher and medium+lower negative by integrating the concept of sentiment valences. The grouping of Psychoticism and Neuroticism instances on the nearest/adjacent plotting points suggested that users from the respective categories of traits might apply similar syntactical and semantical structures of Malay texts to describe their emotions, feelings, and opinions in a negative manner. To the best of the authors' knowledge, no study has evaluated the representation of lower and higher negative words in the Malay Language. From the stance of psychology-personality, particularly Big Five Personality Model, Psychoticism characteristics were viewed as a trait that strongly correlated to Neuroticism and Agreeableness (Van Dam et al., 2005). Although general psychologists have employed factor analysis to determine the relationships between the traits, the strong correlation between Psychoticism and Neuroticism illustrated in plotting of the text instances seems to be in line with that reported by Van Dam et al., (2015). Within the context of text mining, the evidence retrieved from the plotting indicates that FP and FN classification errors may be minimised if personality recognition mechanisms are embedded into the criminality text detection model.

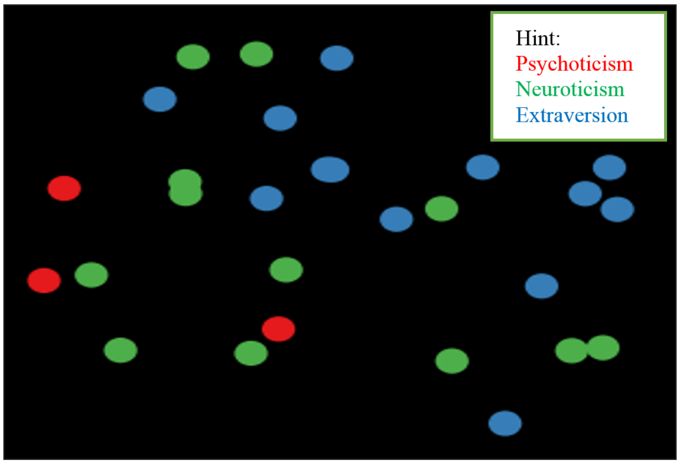


Figure 5. Plotting of PEN model instances

Significant Terms Extraction

Further analysis was performed specifically on Psychoticism instances to seek the nature of writing structures that advocated higher negativity. The Chi Square (X^2) feature identification method was applied to extract the significant terms correlated to Psychoticism. The analysis was conducted from single to triple language models. The highly correlated language models are presented in Table 4, whereby most of the significant terms that featured highly negative/cursing referred to sexuality, sensuality, and sexual expressions. The use of abusive English curse words, such as *fuck* and *shit*, displayed bilingual cursing expressions. Bakar et al., (2018) found that Malaysians prevalently used bilingual or *Bahasa Rojak*, especially a mixture of Malay and English words in social network posts to express their emotions, thoughts, and feelings. Nevertheless, in the context of tweets linked with Psychoticism, the use of profanity English terms became significant and substantially mirrored the latent characteristics of the trait, such as being aggressive and antisocial. Apart from the terms listed in Table 4, other terms, such as *(pu)kimak*, *puki*, *kafir*, *jadah*, and *sial*, were statistically insignificant, but were frequently mentioned by many users.

Table 4
Extracted twitter terms using x^2

Unigram	Bigram	Trigram
Anjing	Gila Babi	Nak Main *Game*
Babi	Anak / Makan Babi	*Fuck Off* la
Tetek	Tetek Kecil	Pakai Kondom tak
Bontot	*Fuck* Orang	*Shit* kalau aku
Melancap	Pakai Kondom	Tetek Size A/kecil
Kote	Lubang Buntut	
Cilaka	Nak Main / *Sex*	
Lancau	Pancut Dalam	
Senggama		

CONCLUSIONS

The role of personality to represent individual characteristic patterns of thoughts, emotions, and behaviour, along with psychological mechanisms, portrays an essential role in every action executed by human. Many researchers have suggested that the human language has rather strong and positive reflections towards their behaviours and intentions, especially within the context of delinquent behaviours. Based on this concept, this study annotated a list of tweets in the Malay language based on a trait that has been frequently used in criminology called Psychoticism by employing three ML algorithms to automatically predict the training instances. To the extent of the authors' knowledge, this appears to be the first study that has measured Malay illicit messages from the stance of personality. The promising accuracy displayed by the algorithms illustrates the ability of automatic classifiers to detect the presence of criminality element in texts. The findings of x^2 based on significant terms exhibited the possibilities of adapting those attributes as complimentary features to improve detection of illicit messages automatically. The analyses outcomes signified that SMO outperformed other classifiers insignificantly by achieving 92.85% of accuracy. Based on x^2, several swear terms, such as *bontot*, *melancap*, and *kote*, displayed significant correlation with Psychoticism Tweets due to the nature of the trait that has been subjected to criminality behaviour, for instance, aggressive and antisocial attributes. The Malay swear words seem more like cursing and taboo, when compared to such words in the English Language. Perhaps, this is due to the curse words in the English language that have been used commonly or the preference amongst the local people in practicing unofficial terms/language (Bahasa Selanga/Pasar), such as *tetek* or *bontot*, in which the standard terms refer to *payu dara* and *punggung* for the respective words. In future, more data will be employed to predict the criminality texts by incorporating the concepts of Artificial Intelligence and Deep Learning.

ACKNOWLEDGEMENT

This study is funded by Universiti Sains Malaysia (USM) RUI grant (1001. PKOMP.8011035).

REFERENCES

Aalderks, D. R. (2014). *Using Latent Semantic Analysis to Detect Non-Cognitive Variables of Academic Performance* (Master Dissertation). New Jersey Institute of Technology, Newark, New Jersey.

Agarwal, B. (2014). Personality detection from text: A review. *International Journal of Computer System, 1*(1), 1-4.

Al-Moslmi, T., Omar, N., Albared, M., & Alshabi, A. (2017). Enhanced Malay sentiment analysis with an ensemble classification machine learning approach. *Journal of Engineering and Applied Sciences, 12*(20), 5226-5232.

Al-Saffar, A., Awang, S., Tao, H., Omar, N., Al-Saiagh, W., & Al-bared, M. (2018). Malay sentiment analysis based on combined classification approaches and Senti-lexicon algorithm. *PLoS One, 13*(4), 1-18.

Allport, G. W., & Odbert, H. S. (1936). Trait-names: A psycho-lexical study. *Psychological Monographs, 47*(1), i-171.

Allport, G. W. (1961). *Pattern and growth in personality*. Oxford, England: Holt, Reinhart & Winston.

Argamon, S., Dhawle, S., Koppel, M., & Pennebaker, J. (2005, July 24-28). Lexical predictors of personality type. In *Proceedings of Interface and the Classification Society of North America* (pp. 1-16), Cincinnati, USA.

Bakar, M. S. A., & Mazzalan, A. M. (2018). Aliran pertuturan bahasa rojak dalam kalangan pengguna facebook di Malaysia [Speech Flow of Mixed Language among Facebook Users in Malaysia]. e-*Academia Journal, 7*(1), 61-71.

Barber, D. (2012). *Bayesian reasoning and machine learning*. Cambridge, UK: Cambridge University Press.

Bertsimas, D., & Dunn, J. (2017). Optimal classification trees. *Machine Learning, 106*(7), 1039-1082.

Bradley, M. M., & Lang, P. J. (1999). *Affective norms for English words (ANEW): Instruction manual and affective ratings* (Vol. 30, No. 1, pp. 25-36). Technical Report C-1, The Centre for Research in Psychophysiology, University of Florida.

Celli, F., Pianesi, F., Stillwell, D., & Kosinski, M. (2013, July 11). Workshop on computational personality recognition: shared task. In *Seventh International AAAI Conference on Weblogs and Social Media* (pp. 1-5). Boston, USA.

Chekima, K., & Alfred, R. (2018). Sentiment analysis of Malay social media text. In R. Alfred, H. Iida, A. A. A. Ibrahim & Y. Lim Eds.), *Computational science and technology* (pp. 205-219). Singapore: Springer.

Correa, T., Hinsley, A. W., & De Zuniga, H. G. (2010). Who interacts on the Web? The intersection of users' personality and social media use. *Computers in Human Behavior, 26*(2), 247-253.

Darwich, M., Noah, S. A. M., & Omar, N. (2016). Automatically generating a sentiment lexicon for the Malay language. *Asia-Pacific Journal of Information Technology and Multimedia, 5*(1), 49-59.

Dunham, M. H. (2006). *Data Mining: Introductory and advance topic*. New Delhi, India: Pearson Education.

Farshad, R., Arefeh, K. S., & Kathayoun, M. (2016). Forensic linguistics in the light of crime investigation. *Pertanika Journal of Social Sciences and Humanities, 24*(1), 375-384.

Gao, Z., Xu, Y., Meng, F., Qi, F., & Lin, Z. (2014, May 11-14). Improved information gain-based feature selection for text categorization. In *2014 4th International Conference on Wireless Communications, Vehicular Technology, Information Theory and Aerospace and Electronic Systems (VITAE)* (pp. 1-5). Aalborg, Denmark.

Garcia, D., Garas, A., & Schweitzer, F. (2012). Positive words carry less information than negative words. *EPJ Data Science, 1*(3), 1-12.

Gerald, M., Ian, J. D. & Martha C. W. (2003). *Personality traits* (2nd Ed.). Cambridge, UK: Cambridge University Press.

Go, A., Bhayani, R., & Huang, L. (2009). Twitter sentiment classification using distant supervision. *CS224N Project Report, Stanford, 1*(12), 1-6.

Goldberg, Y., & Elhadad, M. (2008, June 16-17). splitSVM: fast, space-efficient, non-heuristic, polynomial kernel computation for NLP applications. In *Proceedings of ACL-08: HLT, Short Papers* (pp. 237-240). Columbus, USA.

Guadagno, R. E., Okdie, B. M., & Eno, C. A. (2008). Who blogs? Personality predictors of blogging. *Computers in Human Behavior, 24*(5), 1993-2004.

Hofstee, W. K. (1990). The use of everyday personality language for scientific purposes. *European Journal of Personality, 4*(2), 77-88.

Hossin, M. & Sulaiman, M. N. (2015). A review on evaluation metrics for data classification evaluations. *International Journal of Data Mining and Knowledge Management Process (IJDKP), 5*(2), 1-11.

Huang, F., & Yan, L. (2014). Combined kernel-based BDT-SMO classification of hyperspectral fused images. *The Scientific World Journal, 2014*, 1-13.

Kamaluddin, M. R., Shariff, N. S., Othman, A., Ismail, K. H., & Saat, G. A. M. (2015). Linking psychological traits with criminal behaviour: A review. *ASEAN Journal of Psychiatry, 16*(2), 13-25.

Kapur, B., Ahluwalia, N., & Sathyaraj, R. (2017). Comparative study on marks prediction using data mining and classification algorithms. *International Journal of Advanced Research in Computer Science, 8*(3), 394-402.

Kaur, G., & Chhabra, A. (2014). Improved J48 classification algorithm for the prediction of diabetes. *International Journal of Computer Applications, 98*(22), 13-17.

Korting, T. S. (2006). *C4. 5 Algorithm and Multivariate Decision Trees*. São Paulo, Brazil: National Institute for Space Research–INPE.

Kursuncu, U., Gaur, M., Lokala, U., Thirunarayan, K., Sheth, A., & Arpinar, B. (2018). Predictive analysis on twitter: techniques and applications. In N. Agarwal, N. Dokoohaki & S. Tokdemir (Eds.), *Emerging research challenges and opportunities in computational social network analysis and mining* (pp. 67-104). Cham, Switzerland: Springer.

Mairesse, F., & Walker, M. A. (2011). Controlling user perceptions of linguistic style: Trainable generation of personality traits. *Computational Linguistics, 37*(3), 455-488.

Mitchell, T. M. (1997). *Machine learning*. New York, NY: McGraw Hill Publishers.

Mohammadi, G., & Vinciarelli, A. (2012). Automatic personality perception: Prediction of trait attribution based on prosodic features. *IEEE Transactions on Affective Computing, 3*(3), 273-284.

Nanda, M. A., Seminar, K. B., Nandika, D., & Maddu, A. (2018). A comparison study of kernel functions in the support vector machine and its application for termite detection. *Information, 9*(1), 1-14.

Nasa, C., & Suman, S. (2012). Evaluation of different classification techniques for web data. *International Journal of Computer Applications, 52*(9), 34-40.

Oberlander, J., & Nowson, S. (2006, July 17-21). Whose thumb is it anyway? Classifying author personality from weblog text. In *Proceedings of the COLING/ACL 2006 Main Conference Poster Sessions* (pp. 627-634). Sydney, Australia.

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning*. Urbana, Illinois: University of Illinois Press.

Peng, K. H., Liou, L. H., Chang, C. S., & Lee, D. S. (2015, October 23-24). Predicting personality traits of Chinese users based on Facebook wall posts. In *2015 24th Wireless and Optical Communication Conference (WOCC)* (pp. 9-14). Taipei, Taiwan.

Platt, J. C. (1998). Sequential minimal optimization: A fast algorithm for training support vector machines. *MSRTR: Microsoft Research, 3*(1), 88-95.

Raschka, S. (2014). *Naive bayes and text classification: Introduction and theory*. Ithaca, NY: Cornell university library.

Refaeilzadeh, P., Tang, L., & Liu, H. (2009). Cross-Validation. In L. Liu & Ozsu, M. T. (Eds.), *Encyclopedia of database systems* (pp. 532-538). Boston, MA: Springer.

Rennie, J. D., Shih, L., Teevan, J., & Karger, D. R. (2003). Tackling the poor assumptions of naive bayes text classifiers. In *Proceedings of the 20th International Conference on Machine Learning (ICML-03)* (pp. 616-623). Menlo Park, USA: AAAI Press.

Ruparel, N. H., Shahane, N. M., & Bhamare, D. P. (2013). Learning from small data set to build classification model: A survey. *International Journal of Computer Applications, 975*(8887), 23-26.

Sagadevan, S., Malim, N. H. A. H., & Husin, M. H. (2015). Sentiment valences for automatic personality detection of online social networks users using three factor model. *Procedia Computer Science, 72*, 201-208.

Salloum, S. A., Al-Emran, M., Monem, A. A., & Shaalan, K. (2017). A survey of text mining in social media: facebook and twitter perspectives. *Advances in Science, Technology and Engineering Systems Journal, 2*(1), 127-133.

Saravanan, S. (2016). *Personality Detection in Online Social Networking by using Three Factor Personality Model* (Master dissertation). Universiti Sains Malaysia, Malaysia.

Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Dziurzynski, L., Ramones, S. M., Agrawal, M., ... & Ungar, L. H. (2013). Personality, gender, and age in the language of social media: The open-vocabulary approach. *PloS One, 8*(9), 1-16.

Shally, B. (2014). *Personality Assessment Using Multiple Online Social Networks* (Master dissertation). University of Ottawa, Canada.

Sujatha, R. & Ezhilmaran, D. (2013). Evaluation of classifiers to enhance model selection. *International Journal of Computer Science and Engineering Technology (IJCSET), 4*(1), 16-21.

Urmaliya, A., & Singhai, J. (2013, December 9-11). Sequential minimal optimization for support vector machine with feature selection in breast cancer diagnosis. In *2013 IEEE Second International Conference on Image Information Processing (ICIIP-2013)* (pp. 481-486). Shimla, India.

Van Dam, C., Janssens, J. M., & De Bruyn, E. E. (2005). PEN, Big Five, juvenile delinquency and criminal recidivism. *Personality and Individual Differences, 39*(1), 7-19.

Verhoeven, B., Daelemans, W., & De Smedt, T. (2013, July 11). Ensemble methods for personality recognition. In *Seventh International Association for the Advancement of Artificial Intelligence Conference on Weblogs and Social Media* (pp. 35-38). Boston, USA.

Wang, P., Bojja, N., & Kannan, S. (2015, June 5). A language detection system for short chats in mobile games. In *Proceedings of the third International Workshop on Natural Language Processing for Social Media* (pp. 20-28). Denver, Colorado.

Witten, I. H., Frank, E., & Hall, M. A. (2011). *Data mining: Practical machine learning tools and techniques*. Burlington, MA: Morgan Kaufmann.

Pertanika J. Sci. & Technol. 27 (4): 1821 - 1840 (2019)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Measuring Tajweed Augmented Reality-Based Gamification Learning Model (TARGaLM) Implementation for Children in Tajweed Learning

Nurtihah Mohamed Noor[1]*, Marina Ismail[1], Rahmah Lob Yussof[2] and Fakhrul Hazman Yusoff[1]

[1]*Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia*
[2]*Faculty of Computer and Mathematical Sciences, Cawangan Pahang, Kampus Jengka, 26400 Bandar Tun Abdul Razak Jengka, Pahang, Malaysia*

ABSTRACT

Known for its plain and dry content, tajweed learning is often tedious, particularly for children. Our preliminary study confirmed that most of the learners were uninterested to learn tajweed and this fact was agreed by their educators. Currently, technological advancement helps learning enormously and it has been widely utilised, especially for the digital native generation. Therefore, we highlighted the use of augmented reality and gamification as an attempt to attract children to learn tajweed. Based on experience and motivational theory, tajweed augmented reality-based gamification learning model (TARGaLM) was implemented in the tajweed learning. To investigate the effectiveness of the proposed approach in terms of their emotional engagement (enjoyment) potential and learning performance, 198 children constituted four groups participated in the learning activities. TARGaLM successfully gained positive results of autonomy, challenge, points, badges, leader board, progression, immersion, and feedback, which are crucial for enjoyment. In the post-activity interviews, the students mentioned that the proposed approach was interesting. Furthermore, the proposed approach group indicated the highest decrease in articulation errors from the post-test compared to other groups. The use of current technology and approach not only has potential in bringing enjoyment, but also gives a positive impact on the learning outcomes.

Keywords: ANOVA, augmented reality, enjoyment, experience, gamification, motivation, post hoc test, tajweed learning

ARTICLE INFO

Article history:
Received: 04 April 2019
Accepted: 11 July 2019
Published: 21 October 2019

E-mail addresses:
nurtihah_83@yahoo.com (Nurtihah Mohamed Noor)
marina@tmsk.uitm.edu.my (Marina Ismail)
rahmah128@pahang.uitm.edu.my (Rahmah Lob Yussof)
fakhrul@tmsk.uitm.edu.my (Fakhrul Hazman Yusoff)
* Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

INTRODUCTION

Tajweed knowledge is crucial in every Muslim life as it ensures the correctness of the Quranic recitation. Hence, tajweed learning processes start as early as childhood. Children learning is prominent as Islam emphasises that any religion-related education such as praying, fasting, and Quranic recitation should start at the early age of a Muslim (Ulwan, 2002).

However, previous studies showed that accurate tajweed articulation in Quranic recitations are rather low or moderate for all age levels (Awang et al., 2011; Hassan & Zailaini, 2013a), including the children (Hassan & Zailaini, 2013b; Ismail et al., 2011; Noor et al., 2018). Our preliminary study also found similar result. Shockingly, most of the respondents who were children stated that the tajweed learning processes were tedious and they were uninterested to learn tajweed (Noor et al., 2018). The educators confirmed the findings. Besides, as tajweed has religion-related content, the learning processes are always dry, plain, and limited (Jusoh & Jusoff, 2009).

There are various earlier studies that discussed the use of technological approach to attract learners to learn (Billinghurst et al., 2001; Dunleavy, 2014) the Islamic content specifically (Ismail et al., 2011; Sardan & Rias, 2013; Jusoh & Jusoff, 2009). Augmented reality and gamification are examples of the current technology. Augmented reality enables the virtual contents such as text, animation, sound, and videos to be embedded in the real-world situation (Bakar et al., 2018; Billinghurst, 2002). Gamification permits the utilisation of game elements in the non-game field such as points, level, leader board, and others into the learning process (Barata et al., 2013; Deterding, 2015; Morschheuser et al., 2018). Interactivities, attractiveness, and experiential learning in augmented reality and gamification elicit myriad emotions and appeal to the learners to engage in a learning process (Dunleavy, 2014). Even though there are several studies concerning technological approach to tajweed learning such as Ahsiah et al. (2013), Ismail et al. (2011), and Mssraty and Faryadi (2012), it is still considered scarce compared to other learning content (Noor et al., 2018; Jusoh & Jusoff, 2009). Moreover, through literature search, currently there is still no implementation of augmented reality and gamification employment in tajweed learning, particularly in Malaysia (Noor et al., 2018). The combination of both approaches for learning is possible (Dunleavy, 2014; Noor et al., 2018). However, it requires a thorough design process to avoid cognitive overload during the learning process.

Tajweed learning is often trapped in an uninteresting learning process nature. Therefore, to make the students emotionally engaged in and enjoy the learning process, this study has constructed an innovative model named tajweed augmented reality-based gamification learning model (TARGaLM) as depicted in Figure 1. TARGaLM blends tajweed learning methods (observe, recognise, understand, recall, and articulate), augmented reality, and gamification with four stages of Kolb's experiential learning model, namely concrete experience, reflect observation, abstract conceptualisation, and active experimentation.

Meanwhile, the motivational theory of fantasy, control, curiosity, and challenge was utilised to maximise the learners' engagement in the tajweed learning. The system is coupled with augmented reality and gamification environment to complete the learning processes, which highlight the role of both augmented reality and gamification in enriching experience and interactivity in the learning processes. The system also attracts the learners to learn tajweed, elicits a variety of emotional feelings such as fantasy, control, curiosity, and sense of challenge, as well as improves the learning outcomes.

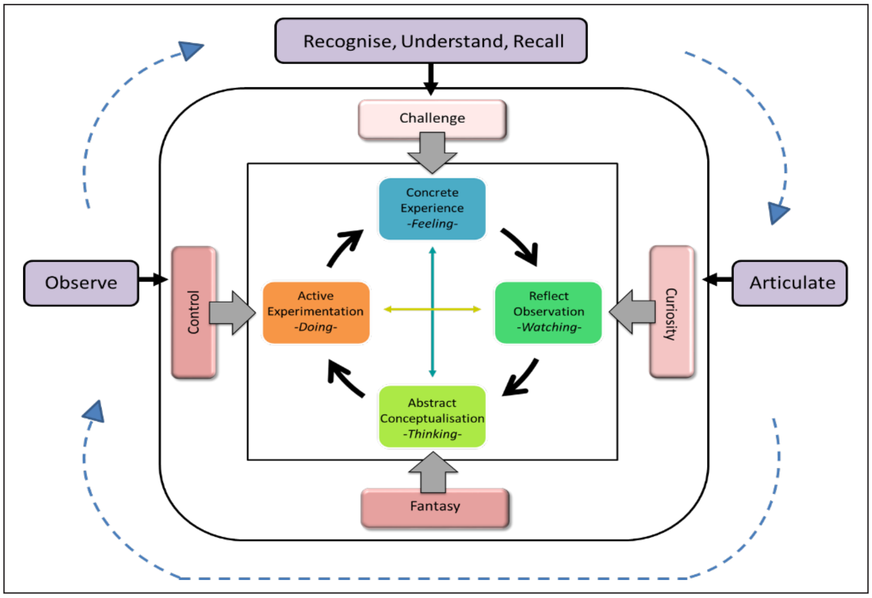


Figure 1. Tajweed augmented reality-based gamification learning model (TARGaLM)

Knowledge gained from experience can definitely be applied to the daily and future life, which is crucial in learning (Huang et al., 2016). Experiential learning is different from the teacher-centred learning as the learners personally control and experience the learning processes (Kolb, 1984). Through interactive learning processes, the learners obtain understanding of the core of the learning content and will parallelly comprehend the relations between the concepts or rules.

Experiential learning model proposed by Kolb argued that learning is a process of experience transformation. The model discussed four cyclic stages of learning, namely concrete experience (CE), reflect observation (RO), abstract conceptualisation (AC), and active experimentation (AE). In the proposed approach, the adapted experiential learning cycles for tajweed learning require:

- CE (feeling) – the learners learn through direct experience and specific experience with peers.
- RO (watching) – the learners observe the experience and make judgements.

- AC (thinking) – the learners learn and are able to understand ideas, concepts, rules, and situations.
- AE (doing) – the learners achieve the goals by completing tasks and activities (challenges).

Motivational theory also plays its own role in ensuring the learners remain in the learning experience (Malone & Lepper, 1987). Fantasy, control, curiosity, and challenge theory of motivation is a robust theory that has been applied in many learning fields (Kim & Lee, 2013), including both augmented reality and game design environment (Dunleavy, 2014). In general, TARGaLM utilises motivational theory in various ways as shown in Table 1.

Experience and motivation adoption in technological learning environment offer one of the utmost emotional effects on the learners, which is enjoyment (Fu et al., 2009; Lumby, 2011; Malone & Lepper, 1987). Concentration, immersion, clear goals, autonomy, feedback, and competence can be used to evaluate enjoyment (Sweetser & Wyeth, 2005) in relation

Table 1
Motivational theory use in TARGaLM

Theory	Description
Fantasy	• Fantasy feeling is evoked in the TARGaLM approach in several ways, for example: • Learn by using augmented reality • Utilisation of animations, audio, or videos, particularly the 3D or 2D animations for explaining the rules of each letter. • Practise the tajweed contents by using a digital game, which emphasised: • The search and select the correct or wrong letters and pronunciations – highlight the use of multimedia elements. • Storyline was formed, particularly in the digital game for completing the learning processes. • Receive rewards • Points, badges, and unlocking contents for learners who complete an activity which evokes the sense of mastery. • Markers' scan in augmented reality • Audio and video utilisation
Control	• Learners can select the markers based on what they want to learn. • Learners can begin the practices by their own selection of topics.
Curiosity	• Curiosity in TARGaLM approach was evoked through (for examples): • Provide the graphics, 2D or 3D graphics, and animations to show the inner mouth for tajweed pronunciations. • Wrong or correct audio recitations of particular Quranic sentences for tajweed. • Myriad levels of challenges in the practices. • Appropriate use of multimedia elements.
Challenge	• Challenge integrated in the overall process (augmented reality for learning, digital game for practising, and audio/video recording for assessment). • The tajweed contents were scaffold based on the complexity (easy, intermediate, expert), particularly in the practice part. • Emphasised the use of game mechanics (points, badge, leader board, and others) for performance feedback and also positive feedback to encourage learners' self-esteem.

to experience and motivation (Fu et al., 2009; Malone & Lepper, 1987). When learners are actively involved in a learning environment such as by dealing with various stimuli from several sources and solve different tasks and activities, the learners' concentration will increase. Fantasy and curiosity feeling elicitation can also attract the learners' attention which help the learners to maintain their focus in the environment (Sweetser & Wyeth, 2005). Besides concentration, well-designed experience and motivation in a learning process is possible to make the learners immerse in the environment (Alexiou et al., 2012). The feeling of less aware of their surrounding area, time, and the current life are examples of immersion effects on the learners.

In any learning environment, clear goals are crucial to ensure the learners will experience activities effectively to achieve the required outcomes (Tan, 2015). Moreover, the sense of fantasy, control, curiosity, and challenge can be properly constructed to motivate the learners when the learning goals and objectives are well-defined (Malone & Lepper, 1987). On the other hand, autonomy has also been proven to be essential in learning experience, particularly to bring enjoyment to the learners (Sweetser & Wyeth, 2005). The processes of feeling, watching, thinking, and doing are significant when the learners are able to control the environment on their own (Lumby, 2011). The experience and motivational theory applied in a learning process are meaningless if the learners cannot retrieve the appropriate feedback in the process (Sweetser & Wyeth, 2005). Feedback covers all the responses provided by the learning processes.

Besides that, experiencing the learning processes and then staying engaged in the environment will also contribute to competence among the learners (Ryan & Deci, 2000). Competence is related to the improvements gained from the learning processes (Fu et al., 2009). Basically, several factors can be used to assess the learners' thoughts regarding their competence such as challenge, knowledge or skill improvement, points, badges, leader board, and progression. To be competent, a learner needs to complete activities or tasks and solve problems (challenges) provided in the learning processes (Iten & Petko, 2014). The improvement of knowledge shows that the learning processes are able to bring competence to the learners. Point and badge collection indicates the effort of the learners in mastering the learning content. Meanwhile, leader board and progression show the achievement or competence levels of the learners compared to others (Codish & Ravid, 2014). Fundamentally, by applying experience and motivational theory in TARGaLM, enjoyment can be achieved successfully by assessing concentration, immersion, clear goals, autonomy, feedback, and competence.

However, a comprehensive implementation of experience and motivation is required to avoid any unattractiveness or cognitive overload of the proposed approach (Dunleavy, 2014; Huang et al., 2016). Accordingly, this study seeks to investigate the effectiveness of the proposed approach for positive emotion, which is enjoyment elicitation potentials and its effect on the performance.

METHOD

To understand the emotional engagement (enjoyment) potential and learning effects of the gamification and augmented reality use in the tajweed learning context, this study adapted enjoyment questionnaires from Fu et al. (2009) and Codish and Ravid (2014) as well as employing pre-test and post-test in the experimental design sessions. The questionnaires consisted of six main dimensions, namely concentration, immersion, clear goals, autonomy, feedback, and competence (knowledge improvement, points, badges, leader board, and progression). Meanwhile, audio or video recording sessions (Hassan & Zailaini, 2013a) were conducted before and after the experimental session as the tajweed learning emphasised the articulation of the letters and verses in Quranic recitation.

A pilot study was conducted to measure the precision of the measures in assessing the emotional engagement. Thirty-three primary school students were involved in the pilot test and a reliability test was performed by computing the Cronbach's coefficient alpha. According to the test results as depicted in Table 2, most of the alpha values were more than 0.7 except for the challenge dimension. Coefficient alpha values above 0.7 were acceptable for the internal consistency of the questionnaires (Nunnally, 1978). However, the constructs in the challenge(*) dimension remained because the values were also acceptable as they were close to 0.7 (Loewenthal, 2004). For the actual experiment, 198 subjects who were the upper-level primary school students were involved in the experimental sessions. The students were randomly assigned into four groups of different learning approaches with approximately equal numbers of male and female students in each group.

Group A was involved in the conventional learning approach (control), Group B only used augmented reality application as well as conventional *practices* and *assessment* (experimental group I), Group C was involved in video learning and gamification

Table 2
Reliability test for the questionnaire

Dimension	Cronbach's Coefficient Alpha values	Number of Items
Concentration	0.790	3
Immersion	0.750	4
Clear Goals	0.887	3
Autonomy	0.749	3
Feedback	0.801	3
Competence:		
Challenge	0.689*	4
Knowledge Improvement	0.712	5
Points	0.767	3
Badges	0.747	5
Leader board	0.854	3
Progression	0.733	4

(experimental group II), and Group D was involved in both the augmented reality and gamification approach (experimental group III) (Huang et al., 2016). The objective of dividing the students into different learning approaches was to investigate the effectiveness of the proposed approach (Group D) compared to the other approaches in tajweed learning in terms of their emotional engagement and performance.

To avoid any interference and invalid data collection, the experimental sessions (briefing, pre-test data collection, and learning activities) and post-test data collection were conducted concurrently for all the groups. Each group was briefed about the experiment procedures. The test sessions began with voice and video recordings of the individual students reciting five surahs from the Quran; Al-Alaq (1-6), Al-Fiil (1-6), Al-Kafirun (1-6), Al-Masadd (1-5), and Al-Ikhlas (1-4) (Hassan & Zailaini, 2013a). These recordings acted as the pre-test data. Upon completion of the recordings, each respective group began their tajweed learning processes. This study utilised the most basic tajweed learning content which were articulation from the vocal cord (*halkum*), through the tongue and nose (nasalisation and non-nasalisation), as well as the Qolqolah rules (Hassan, 2014; Jawatankuasa Buku Teks Sekolah Rendah Agama, 2011). The learning content for each group was identical during the experiments. Figure 2 shows that the tajweed learning requirements are mapped into the process, challenges, and activities in the experiments for each group's learning processes (Kementerian Pendidikan Malaysia, 2017).

Group A is the control group. The group members learnt through the conventional approach. All the learning, practices, and assessments were conducted in a conventional manner. Group B learnt using the augmented reality system only and the *practices* and *assessment* parts were completed conventionally and handled by an educator. Group C used video-based learning and gamification systems and Group D utilised the augmented reality and gamification approach for both the *learning* and *practices.* For augmented

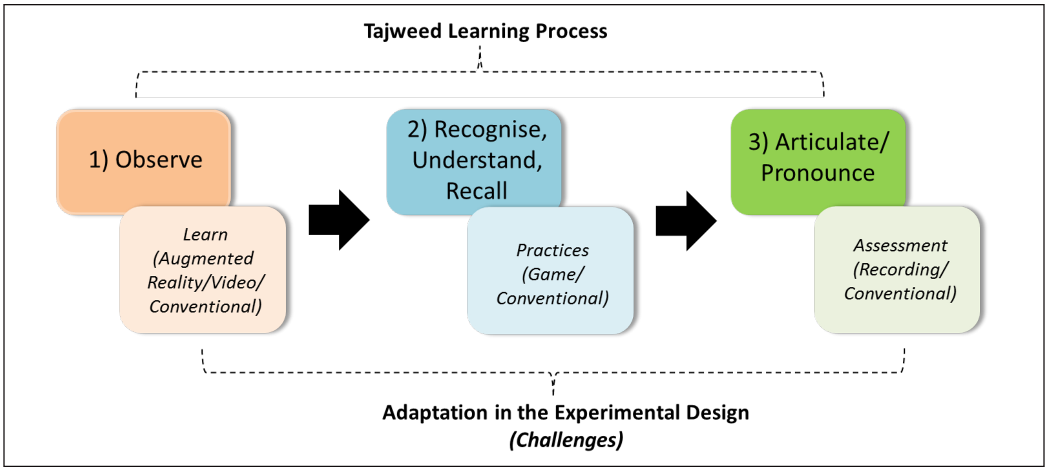


Figure 2. Tajweed learning process and the employment of the approaches

reality system, *hijaiyah* letters were used as the markers as deployed by Rambli et al. (2013). Tajweed learning emphasises articulation of the Quranic letters and verses in certain rules and they were articulated differently (Hassan, 2014). In general, the students in Group D performed the tajweed learning processes as illustrated in Figure 3. This study also adopted crucial elements of gamification employment as described by Tan (2015), which are goal and rules, gameplay, feedback, game space design, and storyline. Further implementation explanations for the TARGaLM application can be retrieved from Noor et al. (2018) in general.

The students in both groups C and D needed to record their recitations individually for the *assessment* part with the help of the researcher. The results of the *assessments* were displayed to the students briefly the next day.

Since students in Group A had friends and peers in completing the learning approach together as mentioned by their educators, students in Groups B, C, and D completed the session in pairs. The decision was made for several reasons. First, as the nature of the tajweed learning emphasises the articulation or pronunciations of the letters or verses, our pilot study revealed that when the students had their peers with them, they tended to repeat the articulation of the learning content simultaneously while completing the learning

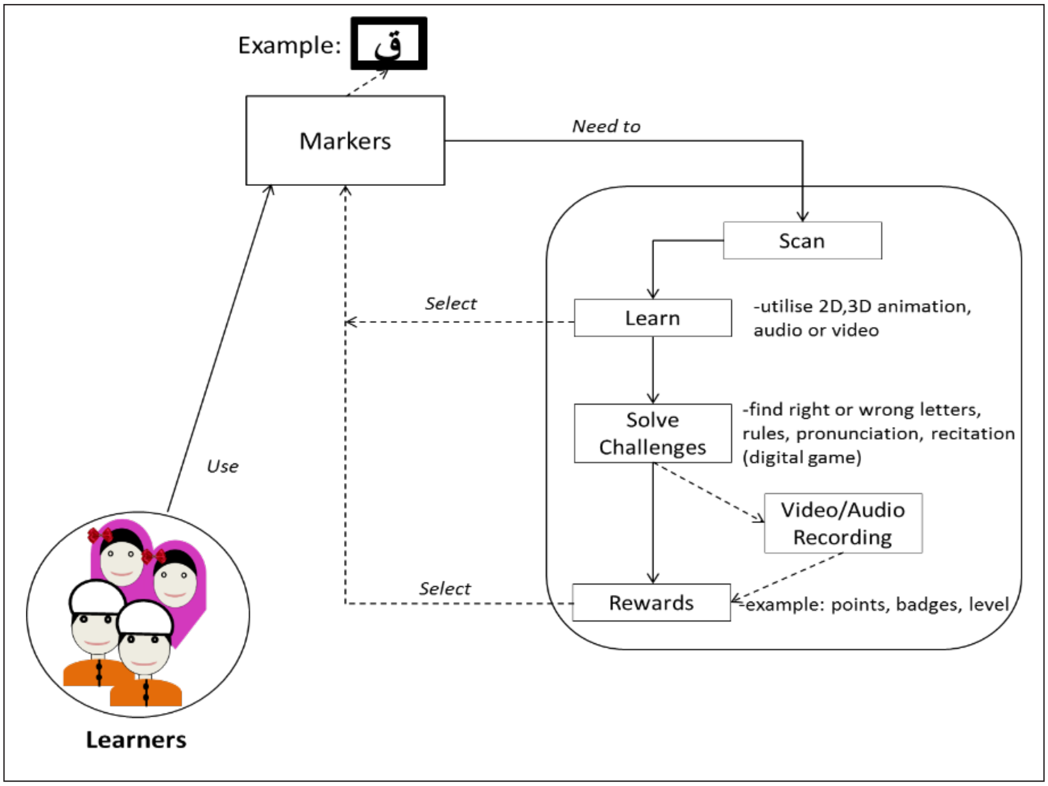


Figure 3. The tajweed learning process for the proposed approach

processes in the approach. However, the researcher asked both of them to interact with the systems equally to ensure they gain experience from the sessions. Second, the students seemed to be more comfortable and relaxed while performing the learning session as he or she had a peer to discuss and solve the tasks and challenges together, which is common in the conventional learning. Furthermore, collaborative learning with peers in a learning session, particularly technological-related approach, enhances the students' confidence and self-efficacy to learn and succeed in the learning process (Sung & Hwang, 2013). It is also an effective opportunity for assuring retention and problem-solving in a learning process (Janssen et al., 2010).

Considering that the students might have a close peer, they were given the options to choose the peer they wanted to be with. After the students finished the learning process, approximately between 75 to 90 minutes, they were asked to complete the post-test. After the post-test recording sessions ended, they answered the provided questionnaire. Table 3 summarises the group division in the study. Meanwhile, Figure 4 depicts the implemented experimental design example in this study.

Table 3
Group division in experimental design

Group	Learning Approach for Experimental Study	Participants	Topics (all groups)
A	Conventional	Pair/Peer	Articulations: - Vocal cord (*halkum*) - tongue - nose (nasalisation and non-nasalisation) - Qolqolah
B	Augmented reality prototype + questions/answers sessions (conventional)	Pair	
C	Video and game prototype (gamification)	Pair	
D	Augmented reality and game prototype (gamification)	Pair	

Figure 4. Experimental design implementation example of the study

RESULTS AND DISCUSSION

This study aims to investigate the effectiveness of the augmented reality and gamification utilisation compared to other approaches in tajweed learning. Evaluations have been deployed to see whether the proposed approach has the potential to elicit positive emotions (enjoyment) as an attempt to eliminate the tedious feeling in tajweed learning and to improve the performance. There are several factors that can be investigated to determine the learners' feelings or emotional attraction towards a technological approach or system (Fu et al., 2009; Malone & Lepper, 1987). This study investigated the factors of concentration, immersion, clear goals, autonomy, feedback, and competence dimension for ensuring the possibility of the positive emotional elicitation in the proposed approach. Those measures were also in line for evaluating learners' experience and motivation in technological environment, which exposed the enjoyment feeling towards the approach (Fu et al., 2009; Malone & Lepper, 1987; Nakamura & Csikszentmihalyi, 2002).

Table 4 shows the demographics of the participants in the experimental study. 198 participants were involved in this evaluation. There was a balanced number of participants for Groups A (51, 25.8%), B (49, 24.7%), C (50, 25.3%), and D (48, 24.2%) as well as for gender; male (106, 53.6%) and female (92, 46.5%). Even though the number of

Table 4
Demographic analysis of the participants (N = 198)

Item	**Classification**	**Frequency (%)**
Group	Group A	51 (25.8)
	Group B	49 (24.7)
	Group C	50 (25.3)
	Group D	48 (24.2)
Gender	Male	106 (53.6)
	Female	92 (46.5)
Age	10	69 (34.8)
	11	64 (32.3)
	12	65 (32.9)
Computer use	Yes	166 (83.8)
	No	32 (16.2)
Playing games on the computer	Often	19 (9.6)
	Rarely	161 (81.3)
	None	18 (9.1)
Augmented reality information	Yes	43 (21.7)
	No	155 (78.3)
Augmented reality use	Yes	35 (17.7)
	No	163 (82.3)
Technology approach for learning	Yes	152 (76.8)
	No	46 (23.2)

participants who often played games on the computer (19, 9.6%) and for both augmented reality information (43, 21.7%) and use (35, 17.7%) is small, the participants were able to use the computer (166, 83.8%), which reasonably indicated their capabilities for involving in the experimental studies.

Table 5 demonstrates tajweed learning proficiency and opinions. The table shows that although the participants were interested in tajweed (107, 54%), their retention (162, 81.8%) and comprehension level (152, 76.8%) were low. In order to perform in the learning, almost half of the participants agreed that the current learning method could be improved (98, 48.5%). The participants noted that technological approach (176, 88.9%) could be used as an alternative for learning tajweed.

Table 5
Tajweed learning proficiency and opinion (N = 198)

Item	Classification	Frequency (%)
Interest	Yes	107 (54)
	Less	91 (46)
	No	0 (0)
Easy to memorise	Yes	26 (13.1)
	Less	162 (81.8)
	No	10 (5.1)
Comprehension level	Yes	34 (17.2)
	Less	152 (76.8)
	No	12 (6.1)
Current learning method preference	Prefer	96 (48.5)
	Can be improved	98 (49.5)
	Does not prefer	4 (2.0)
Technology approach for tajweed learning	Yes	176 (88.9)
	No	22 (11.1)

In order to view the differences from the four groups for the aforementioned measures, ANOVA and post hoc tests were performed. Table 6, Table 7(a), Table 7(b) and Table 7(c) display the results for both analyses.

Table 6
ANOVA test results

	df	SoS	MS	F	Sig.
Concentration					
Between groups	3	0.745	0.248	0.672	0.570
Within groups	194	71.624	0.369		
Immersion					
Between groups	3	5.084	1.695	4.870	0.003*
Within groups	194	67.513	0.348		

Table 6 *(continue)*

	df	SoS	MS	F	Sig.
Clear Goals					
Between groups	3	1.124	0.375	1.031	0.380
Within groups	194	70.515	0.363		
Autonomy					
Between groups	3	35.927	11.976	32.903	0.000*
Within groups	194	70.609	0.364		
Feedback					
Between groups	3	22.292	7.431	18.549	0.000*
Within groups	194	77.713	0.401		
Challenge					
Between groups	3	16.794	5.598	13.760	0.000*
Within groups	194	78.927	0.407		
Knowledge Improvement					
Between groups	3	0.072	0.024	0.089	0.966
Within groups	194	52.361	0.270		
Points					
Between groups	3	19.578	6.526	17.868	0.000*
Within groups	194	70.857	0.365		
Badges					
Between groups	3	49.095	16.365	59.195	0.000*
Within groups	194	53.633	0.276		
Leader board					
Between groups	3	16.903	5.634	10.568	0.000*
Within groups	194	103.431	0.533		
Progression					
Between groups	3	17.882	5.961	18.710	0.000*
Within groups	194	61.805	0.319		

* Significance level p < .05

Table 7(a)
Post hoc test results - multiple comparison

Group		**Mean Difference (I-J) For Each Dimension**			
(I)	**(J)**	**Concentration**	**Clear Goals**	**Autonomy**	**Challenge**
D	A	0.013 (p=1.000)	0.084 (p=0.899)	1.054* (p=0.000)	0.742* (p=0.000)
	B	0.145 (p=0.643)	0.206 (p=0.335)	0.227 (p=0.251)	0.295 (0.107)
	C	-0.003 (p=1.000)	0.051 (p=0.976)	0.066 (p=0.951)	0.073 (p=0.943)

* Significance level p < .05

Table 7(b)
Post hoc test results - multiple comparison

Group		Mean Difference (I-J) For Each Dimension			
(I)	**(J)**	**Knowledge Improvement**	**Points**	**Badges**	**Leader board**
D	A	-0.035 (p=0.987)	0.788* (p=0.000)	0.899* (p=0.000)	0.728* (p=0.009)
	B	-0.019 (p=0.998)	0.251 (p=0.176)	1.060* (p=0.000)	0.332 (p=0.117)
	C	0.014 (p=0.999)	0.051 (p=0.975)	-0.220 (p=0.997)	0.044 (p=0.991)

* Significance level p < .05

Table 7(c)
Post hoc test results - multiple comparison

Group		Mean Difference (I-J) For Each Dimension		
(I)	**(J)**	**Progression**	**Immersion**	**Feedback**
D	A	0.749* (p=0.000)	0.217 (p=0.263)	0.812* (p=0.000)
	B	0.340* (p=0.018)	0.397* (p=0.006)	0.601* (p=0.000)
	C	0.046 (p=0.978)	0.023 (p=0.998)	0.124 (p=0.766)

* Significance level p < .05

From the 11 dimensions investigated, eight dimensions of measurement reached a level of statistical significance ($p < 0.05$). The eight dimensions are autonomy, challenge, points, badges, leader board, progression, immersion, and feedback. Meanwhile, concentration, clear goals, and knowledge improvement showed no significant difference ($p < 0.05$) between the four groups, which indicated that all the groups offered concentration, clear goals, and knowledge improvement to the learners. By looking at the post hoc tests for autonomy, challenge, points, and leader board, the mean difference patterns between Groups D and A were quite similar to each other in which the mean difference is high ($m > 0.7$) for each dimension. However, the mean difference for autonomy was the highest ($m > 1.0$) among those dimensions, which indicated that the learners had intensely less control over the process or content during the conventional learning. Meanwhile, there was only a slightly mean difference ($m < 0.1$) for Group C compared to Group D for those five dimensions. Group C utilised video learning and gamification approach, which could be almost similar with the proposed approach for Group D. Therefore, the learners' perceptions were probably not much different for the dimensions.

Meanwhile, badges, progression, and feedback dimensions indicated that the mean difference patterns between Groups A and B compared to Group D were both high, which were statistically significant ($p < 0.05$). Badge dimension recorded the highest mean difference for both Groups A and B ($m > 0.8$). This indicates that badges have not been utilised in both groups. These results are acceptable as the control group (Group A) and experimental group I (Group B) included conventional learning, practices, and assessment in the processes, which had been reminded to be implemented as similar as the actual and regular class session. For progression and feedback, the mean differences for both dimensions in Group A are the highest ($m > 0.7$). In contrast, even though the mean differences for both dimensions were rather low ($m = 0.340$ and $m = 0.601$) for Group B as the values did not exceed $m > 0.7$, the significant values ($p = 0.018$ and $p = 0.000$) showed that they were statistically significant compared to Group D. The learners in Group A perhaps had a confusion about how they saw their progression in the conventional approach. Even though we believed that all groups received appropriate feedback from educators or the developed system, the mean difference ($m > 0.6$) showed that feedback needed were fewer in both Groups A and B compared to Group D. In general, some of the results related to points, badges, leader board, and progression were in line with Codish and Ravid's (2014) study. Even though the dimensions were significant for the groups, there were various mean value differences between the dimensions which showed that the learners' view could be differentiated by the learning approach and types of learners. Since the participants were randomly selected, the various types of learners might blend in the groups of the study.

The mean differences for concentration, clear goals, and knowledge improvement dimensions between Groups A, B, and C compared to Group D were relatively small ($m > 0.7$) and non-significant ($p > 0.05$). Some of the results showed that the mean values were negative, which indicated the learners in the comparison group favoured the dimension over Group D. There was a negative mean difference for concentration between Groups D and C. However, the value was only slightly different ($m = -0.003$). This result was supported by Fu et al. (2009), who clarified that a more complex technological learning approach might reflect the learners' preferences and focus.

For knowledge improvement, the results showed that there were negative mean differences for both Groups A and B ($m = -0.035$ and $m = -0.019$) compared to Group D, with only a slight difference for Group B. Although previous studies found that technological learning approach could improve knowledge (Fu et al., 2009; Sweetser & Wyeth, 2005), the conventional approach was still relevant to improve the learners' knowledge (Codish & Ravid, 2014; Sabu, 2012). In addition, in this study, the conventional learning routine of tajweed learning might make some of the learners in Group A believed that the conventional learning approach was the most effective for their knowledge improvement. Furthermore, the utilisation of new and less exposed technology or approach to tajweed learning might invite doubts to its effectiveness for knowledge improvement and retention. Compared to Group D, only Group B shows that there was a significant difference in immersion dimension, while Groups A and C were not significant. The utilisation of augmented reality only in Group B learning processes might be one of the reasons for the children feeling less immersed.

On the other hand, we also collected pre-test and post-test data of Quranic recitations before and after the learning processes to investigate the performance of the proposed approach compared to others. Descriptive and paired samples *t* test analyses were applied to analyse 360 audio or video data from 180 participants, after excluding several missing and corrupted files. Table 8 and Table 9 show the results.

Table 8
Paired samples statistics and tests for audio/video recordings of Quranic recitations

G.	**Test**	**N**	**Mean**	**SD**	**t**	**df**	**Sig.**
A	Pre	40	16.1000	6.38026	2.317	39	0.026*
	Post	40	14.5750	5.24716			
B	Pre	46	15.2174	8.02472	2.564	45	0.014*
	Post	46	13.8478	6.63984			
C	Pre	47	12.5319	6.58011	1.521	46	0.135
	Post	47	11.5532	5.94470			
D	Pre	47	15.2553	6.48901	4.195	46	0.000*
	Post	47	12.7234	5.91131			

* Significance level $p < .05$

Contrary to Group C, the results indicated that there was a significant difference for three groups, which were Groups A, B, and D with the values of $p = 0.026$, $p = 0.014$ and $p = 0.000$ in the Quranic recitation. For the proposed approach, Group D indicated the most significant value ($p = 0.000$) whereby the mean values exhibited the highest reduction of Quranic articulation errors from pre-test ($m = 15.2553$) to post-test (12.7234). These findings were supported by previous studies (Ismail et al., 2011; Muhammad et al., 2012), whereby the tajweed knowledge and articulations were improved after utilising technological learning approaches for game and intelligent application. Tajweed learning is known for its dry, plain, and limited learning approach, which makes the learners feel bored and uninterested (Noor et al., 2018; Jusoh & Jusoff, 2009). Augmented reality and gamification make the learning alive and give positive outcomes to the performance. Even though the learners had little doubt on the proposed approach in the knowledge improvement dimension, the results proved that the tajweed knowledge and use improved well from the learning approach process. However, the results showed that the learning performance was not significant for Group C, which meant the articulation errors did not significantly decrease after the learning process. Group C completed video-based learning and gamification approach in the learning process. Apparently, the result might be due to several conditions, which are (a) the learners' capabilities in digesting the learning content during the learning processes and (b) the various types of learners' learning styles (Gagné & White, 1978) could also be a factor, which referred to learners who were not technologically oriented to learn.

Table 9
Audio/video recording number of articulation errors based on surah and categories for all groups

Surah	Test	Articulation Category			
		Vocal Cord	Tongue	Nose	Qolqolah
Al-Ikhlas	Pre	n/a	3	n/a	483
	Post	n/a	0	n/a	404
Al-Masad	Pre	37	25	71	626
	Post	38	27	38	578
Al-Kafirun	Pre	30	65	300	225
	Post	12	32	308	230
Al-Fiil	Pre	7	137	249	15
	Post	4	86	279	7
Al-Alaq	Pre	8	3	39	343
	Post	3	2	19	332

n/a = not applicable

To describe further on performance, Table 9 shows the results of the articulation error categories (vocal cord, tongue, nose, and Qolqolah) based on the Surah recitation. The table indicates that most articulation errors in the post-test for each category decreased, except for Al-Masad, Al-Kafirun, and Al-Fiil. Al-Masad and Al-Kafirun encountered a slight increase of errors in the post-test results for vocal cord, tongue, and nose articulation. However, it seems that Al-Fiil has quite a high increase in the errors for nose articulation. Since nose articulation is related to nasalisation and non-nasalisation rules, the increase is acceptable since the rules are one of the many errors made by the learners (Hassan, 2014) and sometimes the learners may be confused about when to apply the nasal or non-nasal recitations of the Quranic verses. The findings of the articulation errors were in line with past studies (Awang et al., 2011; Hassan & Zailaini, 2013a, 2013b; Ismail & Zakaria, 2009; Basah, 2010; Kamarudin & Salam, 2012). In fact, when a learner has the knowledge of the tajweed rules and on how the rules are articulated while reciting, the learner's articulations of the rules in the Quranic recitation will be better (Hassan, 2014; Hassan & Zailaini, 2013a, 2013b).

In general, the proposed approach (Group D) potentially evoked the learners' positive emotions in terms of autonomy, challenges, points, and leader board over the conventional approach (Group A). It provides the learners more immersion feeling compared to Group B. Besides, badges, progression, and feedback positively existed in Group D compared to Groups A and B, which helped the learners to have feelings of competence, awareness, and guidance. Even though there were no significant differences for concentration, clear goals, and knowledge improvement among the groups, the factors still existed in the proposed approach. Those factors ensured that experience and motivation were employed appropriately in the proposed approach, which could also be factors to avoid tediousness in the tajweed learning process. Furthermore, at the end of the sessions, learners in Group D stated that *"The learning process is very interesting"*, *"Can I use the system again tomorrow?"*, *"Wow, I never used this type of approach before, it is so interesting"*. Meanwhile, the learners in other groups mentioned, *"I want to involve in the session too"*, *"That session seems very interesting"*, *"I need to try playing with the cards too"* and many more. The statements provide additional proofs that augmented reality and gamification utilisation triggered the learners' interest and are capable of eliciting positive emotions for the learners to learn.

In terms of performance, the proposed learning approach also shows the most positive outcomes in minimising errors in the Quranic recitation. This describes the virtual-actual interaction created by the approach as well as the availability of appropriate experience, challenge, curiosity, and fantasy can be a factor in facilitating knowledge retention.

CONCLUSION AND SUGGESTION

This study employs augmented reality and gamification approach for tajweed learning in order to minimise tedious feeling towards a dry, plain, and unattractive learning content, and then attract the children to learn tajweed. The results show that the proposed approach gained positive feedback in terms of autonomy, challenge, points, badges, leader board, progression, immersion, and feedback measures of the study. Positive results of the measures exhibit that the learners feel the experience and motivation better than other approaches. Even though concentration, clear, goals and knowledge improvement do not have significant difference among the groups, it still shows that the factors existed in the proposed approach as well as the other groups. Furthermore, positive improvement of Quranic recitations also indicates that the elicitation of positive emotions through experience and motivation facilitates the cognitive process and retention of the learners, which are consistent with past studies by Ismail et al. (2011) and Muhammad et al. (2012).

In general, this study focuses on the augmented reality and gamification approach for tajweed learning. Future work can focus on other Islamic content learning and different implementations since the approach can be implemented in myriad ways such as location-based augmented reality (Dunleavy, 2014), developing the learners' own content (Tan, 2015), and many more (Perry, 2015; Villagrasa et al., 2014). Longitudinal study with different measurement of investigations (Huang et al., 2016) can also enrich the augmented reality and gamification employment in the learning environment.

ACKNOWLEDGEMENTS

This research received the United States Fulbright-MCMC Senior Specialist visit grant for 2016. The authors fully acknowledged the Malaysian-American Educational Exchange and The Malaysian Communications and Multimedia Commission for approving the fund which made the research guided and feasible in its early phases.

REFERENCES

Ahsiah, I., Noor, N. M., & Idris, M. Y. I. (2013, September 28-29). Tajweed checking system to support recitation. In *2013 International Conference on Advanced Computer Science and Information Systems (ICACSIS)* (pp. 189-193). Bali, Indonesia.

Alexiou, A., Schippers, M., & Oshri, I. (2012). Positive psychology and digital games: The role of emotions and psychological flow in serious games development. *Psychology*, *3*(12), 1243-1247.

Awang, A., Mat, A. C., & Yaakub, A. N. (2011). Tahap pembacaan Al-Quran dalam kalangan pelajar di UiTM Terengganu. *Asean Journal of Teaching and Learning in Higher Education*, *3*(2), 83-100.

Bakar, J. A. A., Gopalan, V., Zulkifli, A. N., & Alwi, A. (2018). Design and development of mobile augmented reality for physics experiment. *Communications in Computer and Information Science, 886*, 47-58.

Barata, G., Gama, S., Jorge, J., & Gonçalves, D. (2013, September 11-13). Engaging engineering students with gamification: An empirical study. In *5th International Conference on Games and Virtual Worlds for Serious Applications (VS-GAMES)* (pp. 1-8). Poole, UK.

Basah, I. K. (2010). *Kemahiran bacaan Al-Fatihah dalam kalangan jemaah masjid: Kajian di masjid-masjid daerah Hulu Langat, Selangor* (Master's thesis). Universiti Malaya, Kuala Lumpur.

Billinghurst, M. (2002). *Augmented reality in education*. Seattle, WA: New Horizons for Learning - Technology in Education. Retrieved July 16, 2014, from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.526.1919&rep=rep1&type=pdf

Billinghurst, M., Kato, H., & Poupyrev, I. (2001). The MagicBook - Moving seamlessly between reality and virtuality. *Computer Graphics and Applications, IEEE*, *21*(3), 6-8.

Codish, D., & Ravid, G. (2014). Academic course gamification: The art of perceived playfulness. *Interdisciplinary Journal of E-Learning and Learning Objects*, *10*(1), 131-151.

Deterding, S. (2015). The lens of intrinsic skill atoms: A method for gameful design. *Human-Computer Interaction - Special Issue "HCI and Digital Games," 30*(3-4), 294-335.

Dunleavy, B. M. (2014). Design principles for augmented reality learning. *TechTrends*, *58*(1), 28-35.

Fu, F. L., Su, R. C., & Yu, S. C. (2009). EGameFlow: A scale to measure learners' enjoyment of e-learning games. *Computers and Education*, *52*(1), 101-112.

Gagné, R. M., & White, R. T. (1978). Memory structures and learning outcomes. *Review of Educational Research*, *48*(2), 187-222.

Hassan, S. S. (2014). *TALAQQI - Asas sebutan huruf Al-Quran* (3rd ed.). Sungai Buloh, Malaysia: Kelas Bimbingan Rehal.

Hassan, S. S., & Zailaini, M. A. (2013a). Analysis of tajweed errors in Quranic recitation. *Procedia - Social and Behavioral Sciences*, *103*, 136-145.

Hassan, S. S., & Zailaini, M. A. (2013b). Khatam Al-Quran in Islamic education curriculum in Malaysia. *Procedia - Social and Behavioral Sciences*, *103*, 558-566.

Huang, T. C., Chen, C. C., & Chou, Y. W. (2016). Animating eco-education: To see, feel, and discover in an augmented reality-based experiential learning environment. *Computers and Education*, *96*, 72-82.

Ismail, M., Diah, N. M., Ahmad, S., & Rahman, A. A. (2011, December 14). Engaging learners to learn tajweed through active participation in a multimedia application (TaLA). In *Proceedings of The 3rd International Conference on Advances in Computing, Control and Telecommunication Technologies* (pp. 88-91). Jakarta, Indonesia.

Ismail, M. F., & Zakaria, S. N. (2009). Kajian pembacaan surah Al-Fatihah di kalangan pelajar Politeknik Kuching Sarawak. In *Seminar Cabaran Pendidikan Di Politeknik*. Kuching, Sarawak. Retrieved July 16, 2014, from http://dspace.poliku.edu.my/xmlui/bitstream/handle/123456789/319/Kajian Pembacaan surah al-fatihah.pdf?sequence=1&isAllowed=y

Iten, N., & Petko, D. (2014). Learning with serious games: Is fun playing the game a predictor of learning success? *British Journal of Educational Technology*, *47*(1), 151-163.

Janssen, J., Kirschner, F., Erkens, G., Kirschner, P. A., & Paas, F. (2010). Making the black box of collaborative learning transparent: Combining process-oriented and cognitive load approaches. *Educational Psychology Review*, *22*(2), 139-154.

Jawatankuasa Buku Teks Sekolah Rendah Agama. (2011). *Tajwid*. Selangor, Malaysia: Jabatan Agama Islam Negeri.

Jusoh, W. N. H. W., & Jusoff, K. (2009). Using multimedia in teaching Islamic Studies. *Journal of Media and Communication Studies*, *1*(5), 86-94.

Kamarudin, N., & Salam, S. (2012). Tajweed understanding among Malaysian Muslims (preliminary findings). In *Proceedings of the 1st International Conference on Mobile Learning, Applications and Services (Mobilecase2010)*. (pp. 1-4). Kuala Lumpur, Malaysia.

Kementerian Pendidikan Malaysia. (2017). *Rancangan pengajaran tahunan (KSSR) Pendidikan Agama Islam*. Putrajaya, Malaysia: Kementerian Pendidikan Malaysia.

Kim, J. T., & Lee, W. H. (2013). Dynamical model for gamification of learning (DGML). *Multimedia Tools and Applications*, *74*(19), 8483-8493.

Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development*. New Jersey, USA: Prentice Hall.

Loewenthal, K. M. (2004). *An introduction to psychological tests and scales* (2nd ed.). London, England: Psychology Press.

Lumby, J. (2011). Enjoyment and learning: Policy and secondary school learners' experience in England. *British Educational Research Journal*, *37*(2), 247-264.

Malone, T. W., & Lepper, M. R. (1987). Making learning fun: A taxonomy of intrinsic motivations for learning. In R. E. Snow & M. J. Farr (Eds.), *Aptitude, learning and instruction volume 3: Conative and affective process analyses* (pp. 223-253). New Jersey, USA: Lawrence Erlbaum Associates, Publishers.

Morschheuser, B., Hassan, L., Werder, K., & Hamari, J. (2018). How to design gamification? A method for engineering gamified software. *Information and Software Technology*, *95*, 219-237.

Mssraty, T., & Faryadi, Q. (2012). Teaching the Quranic recitation with harakatt: A multimedia-based interactive learning method. *International Journal of Scientific and Engineering Research*, *3*(8), 1-4.

Muhammad, A., Ul Qayyum, Z., Mirza, W., Tanveer, S., Martinez-Enrique, A., & Syed, A. Z. (2012). E-Hafiz: intelligent system to help Muslims in recitation and memorization of Quran. *Life Science Journal*, *9*(1), 534-541.

Nakamura, J., & Csikszentmihalyi, M. (2002). The concept of flow. In C. R. Snyder & S. J. Lopez (Eds.). *Handbook of positive psychology* (pp. 89-105). New York, NY: Oxford University Press.

Noor, N. M., Yussof, R. L., Yusoff, F. H., & Ismail, M. (2018). Gamification and augmented reality utilization for Islamic content learning: The design and development of tajweed learning. *Communications in Computer and Information Science*, *886*, 163-174.

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York, NY: McGraw-Hill.

Perry, B. (2015). Gamifying French language learning: A case study examining a quest-based, augmented reality mobile learning-tool. *Procedia - Social and Behavioral Sciences*, *174*, 2308-2315.

Rambli, D. R. A., Matcha, W., & Sulaiman, S. (2013). Fun learning with AR alphabet book for preschool children. *Procedia Computer Science*, *25*, 211-219.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68-78.

Sabu, S. (2012). *Educational psychology*. New Delhi, India: APH Publishing Corporation.

Sardan, N. A., & Rias, R. M. (2013). M-Tajweed: A mobile courseware to assist in tajweed learning. *International Journal on Islamic Applications in Computer Science And Technology*, *1*(3), 1-8.

Sung, H. Y., & Hwang, G. J. (2013). A collaborative game-based learning approach to improving students' learning performance in science courses. *Computers and Education*, *63*(2013), 43-51.

Sweetser, P., & Wyeth, P. (2005). GameFlow: A model for evaluating player enjoyment in games. *ACM Computers in Entertainment (CIE)*, *3*(3), 3-3.

Tan, W. H. (2015). *Gamifikasi dalam pendidikan*. Perak, Malaysia: Universiti Pendidikan Sultan Idris.

Ulwan, A. N. (2002). *Pendidikan anak dalam Islam* (9th ed., vol. 2). Singapore: Pustaka Nasional Pte Ltd.

Villagrasa, S., Fonseca, D., & Durán, J. (2014, October 1-3). Teaching case: Applying gamification techniques and virtual reality for learning building engineering 3D arts. In *Proceedings of the 2nd International Conference on Technological Ecosystems for Enhancing Multiculturality* (pp. 171-177). Salamanca, Spain.

Pertanika J. Sci. & Technol. 27 (4): 1841 - 1862 (2019)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Review article

Sago Wastes and Its Applications

Nurhusni Amin[1], Nordin Sabli[1, 2]*, Shamsul Izhar[1] and Hiroyuki Yoshida[1]

[1]*Department of Chemical and Environmental Engineering, Faculty of Engineering, Universiti Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia*
[2]*Instititute of Advance Technology (ITMA), Universiti Putra Malaysia, 43400, UPM Serdang, Selangor, Malaysia*

ABSTRACT

The sago starch industry is one of the major revenue sources of the Malaysian state of Sarawak. This state is currently among the world's leading producers of sago starch, exporting more than 40,000 tons every year to different Asian countries. This number is expected to rise since starch production and export value have been increasing 15.0%–20.0% each year. Sago palm is subjected to various processes to obtain starch from its trunk. During processing, a huge amount of residual solid wastes is generated, such as bark and *hampas*, and in general, is burned or washed off to nearby streams. Along with the rising sago starch demand, the sago starch industry is now facing waste management problems, which have resulted in environmental pollution and health hazards. These wastes comprise starch, hemicellulose, cellulose, and lignin; hence, can be valorized into feedstock as value-added products. To date, these wastes have been utilized in the production of many materials like adsorbents, sugars, biofuels, nanomaterials, composites, and ceramics. This review article aims to summarize the various methods by which these wastes can be utilized besides to enlighten the major interest on sago *hampas* and bark.

Keywords: Adsorbent, biofuel, sago bark, sago *hampas*, waste utilization

ARTICLE INFO

Article history:
Received: 06 September 2018
Accepted: 23 April 2019
Published: 21 October 2019

E-mail addresses:
nurhusni_amin@yahoo.com (Nurhusni Amin)
nordin_sab@upm.edu.my (Nordin Sabli)
shamizhar@upm.edu.my (Shamsul Izhar Siajam)
h-yoshida@upm.edu.my (Hiroyuki Yoshida)
* Corresponding author

INTRODUCTION

Sago is a Javanese term that means starch-containing palm pith. Scientifically, it is known as *Metroxylon sago,* whereby *metra* means parenchyma or pith and *xylon* means xylem (Singhal et al., 2008). Sago palm is a starch crop par excellence, as it has a higher starch production capacity, between 20–25

ISSN: 0128-7680
e-ISSN: 2231-8526

tons/ha/year, than that of cassava, rice or corn (Flores, 2008). In Malaysia, Sarawak has the largest sago palm plantation areas, particularly in the Mukah division. Currently, it is known as one of the world's largest exporter of sago starch, exporting more than 40,000 tons annually to different countries, including to Peninsular Malaysia. Nevertheless, this value is expected to rise every year, corresponding to the world demand and will subsequently increase the amount of waste generated. The three main types of wastes from the sago industry are sago *hampas*, bark, and wastewater. These wastes predominantly comprise starch and lignocellulosic compounds, namely, lignin, cellulose, and hemicellulose. Sago wastes have been neglected despite its high value. The wastes are not fully utilized as higher value-added products. Moreover, current improper disposal practices of these wastes, such as washed off into nearby rivers, pose harmful effects to the environment.

Waste lignocellulosic biomass is typically an abundant nonedible plant material that has been discarded across the world. This carbon neutral material is generated naturally from available carbon dioxide, water, and sunlight via the photosynthesis process. Therefore, it is considered the only sustainable source of organic carbon and the perfect equivalent of fossil resources. Unlike lignocellulosic carbons, the high dependency of industrial chemicals and synthetic polymers on depleting fossil resources will cause a lethal threat to human beings (Isikgor & Becer, 2015). Hence, elevating concerns regarding this issue has turned researchers' attention to valorize the starchy-lignocellulosic compounds present in sago wastes. Many studies have been carried out to investigate the conversion of this waste into useful products or its use as a raw material in other industries. Wastes produced from the sago industry can be utilized in the production of adsorbents, sugars, biofuels, nanomaterials, composites, and ceramics. Therefore, in this present study, efforts have been made to compile and discuss the recent methods employed to utilize solid wastes from the sago industry, i.e. *hampas* and bark. Besides that, this review will also elucidate the major interest in both wastes.

SAGO PALM

Sago palm belongs to the genus *Metroxylon*, under the order Arecales, family Palmae, and subfamily Calamoidae (Awg-Adeni et al., 2010). It is a hapaxanthic or monocarpic plant that flowers only once in a lifetime and dies shortly thereafter (Bujang, 2006; Flach & Schuiling, 1988). The plant is believed to convert all of its stored nutrients into starch and fill its trunk during the vegetative stage just before flowering (Abd-Aziz, 2002). Sago palm's vegetative phase lasts 7 to 15 years, during which the pith will be saturated with starch, from the base of the stem upwards (Karim et al., 2008). This plant is extremely hardy and is the only starch crop that thrives and grows with minimum care in swampy, acidic peat soils with a low level of nutrients, where very few crops can survive (Flach & Schuiling, 1988; Singhal et al., 2008).

Generally, the sago plant reproduces via suckers that emerge from roots or lower trunks of parent plants, unlike other plants of the same genus that usually propagate by seeds. Nonetheless, propagation from suckers is usually limited in number, hence, an in vitro technique is most suitable to produce a vast amount of materials for extensive planting and to improve the quality and vigor of palms (Singhal et al., 2008). Apart from suckers, sago palm may also propagate from seedlings (Karim et al., 2008). A full-grown plant can grow up to 25.0 m tall, and typically reach a height of around 10.0–12.0 m and a diameter of about 0.800–1.00 m when matured (Bujang, 2006). A matured sago produces the highest starch content of 25 tons per hectare, compared to other starch crops like rice, corn, wheat, and potato, which only produces 6.00, 5.50, 5.00, and 2.50 tons, respectively (Awg-Adeni et al., 2010).

There are several remarkable environmental benefits of sago palm highlighted in previous studies. Firstly, its large fibrous root system helps to trap silt loads and removes pollutants, fecal contaminants, and heavy metals. This feature also makes the plant resistant to floods, drought fires, and strong winds. Next, the sago forest acts as an excellent carbon sink for carbon sequestration (Chew et al., 1999), by assimilating carbon dioxide for starch conversion. Thus, it is a very good tool for offsetting the greenhouse effect and a potential plant resource for bioremediation (Flores, 2008).

The sago palm flourishes in a moist environment and is therefore found along riverbanks, water holes, and swampy areas. It is an indigenous species of the Southeast Asia region, specifically Malaysia, Indonesia, Papua New Guinea, and the Philippines. As such, there are similarities in the terminologies used. In these countries, sago is either known as *saksak*, *rumia*, *lumbia*, or *unod* or the food ingredient *landang* or *kinugay* (Flores, 2008; Naim et al., 2016). Besides that, the sago palm's center of diversity is in the island of New Guinea.

SAGO INDUSTRY

Papua New Guinea, Indonesia, and Malaysia are the three leading producers of sago, whereby sago is primarily grown for sago starch. In Malaysia, more than 90.0% of all sago-planting areas are in the state of Sarawak, which is one of the world's largest sago exporter, exporting about 25,000–40,000 tons of sago products annually to Peninsular Malaysia, Japan, Taiwan and Singapore (Bujang, 2006; Singhal et al., 2008). Currently, Malaysia is the third main sago producer in the world after Indonesia and Papua New Guinea and combined, the three countries produce approximately 94.6% of the world's sago (Naim et al., 2016). Although Indonesia and Papua New Guinea produce more sago than Malaysia, the sago industry in this country is well established and has become one of the important industries contributing to export revenue (Karim et al., 2008). In Malaysia, sago starch ranks fifth in terms of highest agricultural revenue after pepper, palm oil, cocoa, and rubber (Abd-Aziz, 2002).

Sago is cultivated in wild and semi-wild conditions in Sarawak and covers approximately 19,702 ha. The production capacity of sago palm varies from 2.00–5.00 tons of dry starch/ha in the wild to 10.0–25.0 tons/ha in the case of cultivated plants (Abd-Aziz, 2002). Realizing sago's potential, the Sarawak government appointed the Land Custody and Development Authority (LCDA) to kick-start the commercial cultivation of sago and implement management techniques. A large-scale sago palm plantation was initiated in 1986, with the opening of three precedent plantations, namely Mukah Sago Plantation (MSP), Sebakong Sago Plantation (SSP), and Dalat Sago Plantation (Aziin & Rahman, 2005). In fact, Malaysia is the first country that introduced estate plantation for sago (Naim et al., 2016). Based on the latest data by Department of Agriculture Sarawak (DAS), (Department of Agriculture Sarawak, 2013) the total estimated area of sago plantation in 2013 was 54,087 ha, with the largest area found in Mukah division with 46,924 ha. Although estate plantations have been introduced, small-holdings still amount to a huge area, i.e. 37,028 ha.

SAGO STARCH

Processing of Sago Starch

Sago palms are usually felled about 10 to 15 years after planting, depending on the fertility of the terrain. Due to the rapid decline of starch content in the trunk after fruiting, palms at a height between 7.50 to 9.00 m will be felled right after flowering and immediately before fruiting (Singhal et al., 2008). The sago trunk is then cut into about 10–12 sections, with each section measuring 75.0 to 90.0 cm long. Cut sago logs are tied together in the form of rafts and tugged by boats to sago factories for processing. Otherwise, the logs will be kept in ponds prior to processing to reduce deterioration (Vikineswary et al., 1994). Other than that, the logs may also be delivered by land to the factories, in which independent agents will collect and transport the logs for a fee. Since sago factories are invariably situated along river banks, floating of sago logs along the canals and rivers is the most typical form of transportation. However, when the canals become dry, lorries will be used instead (Chew et al., 1999; Chew et al., 1998).

The sago palm pith mainly consists of starch, which has to be separated from the cellulosic materials and thus, will undergo several stages of processing to extract a good quality and quantity of starch (Awg-Adeni et al., 2010). Generally, sago starch extraction can be categorized into two types, traditional and modern. For the traditional method, it is further classified into two levels, namely, domestic and small-scale processing plant. Unlike the small-scale processing plant level, the domestic level is practiced by individual farmers, in which sago palms are felled and processed in the garden without the need to transport the heavy trunks to factories. On the other hand, for modern method, several modifications to the small-scale processing plant level are involved. Mostly, large-scale factories in Sarawak have adopted this technique whereby they are fully mechanized and capable to reduce most laborious tasks (Karim et al., 2008). Nevertheless, the principles and methods of starch extraction are similar in both traditional and modern methods, differing only in scale of operation (Kamal et al., 2007)

During processing, the bark-like layer of the palm is first stripped off followed by maceration using a rasper. The rasping process is done either manually or mechanically to pulverize the pith and loosen the starch particles within the fiber. Next, the starch is removed from the fiber either by kneading with hand, trampling by feet or spraying with water. Then, the starch is passed through a series of centrifugal sieves for removal of coarse fibers. It is later extracted using cyclone separators and then dried with a rotary drum dryer, followed by hot air drying (Chew et al., 1998; Kamal et al., 2007; Vikineswary et al., 1994). A simplified sago starch extraction process is shown in Figure 1.

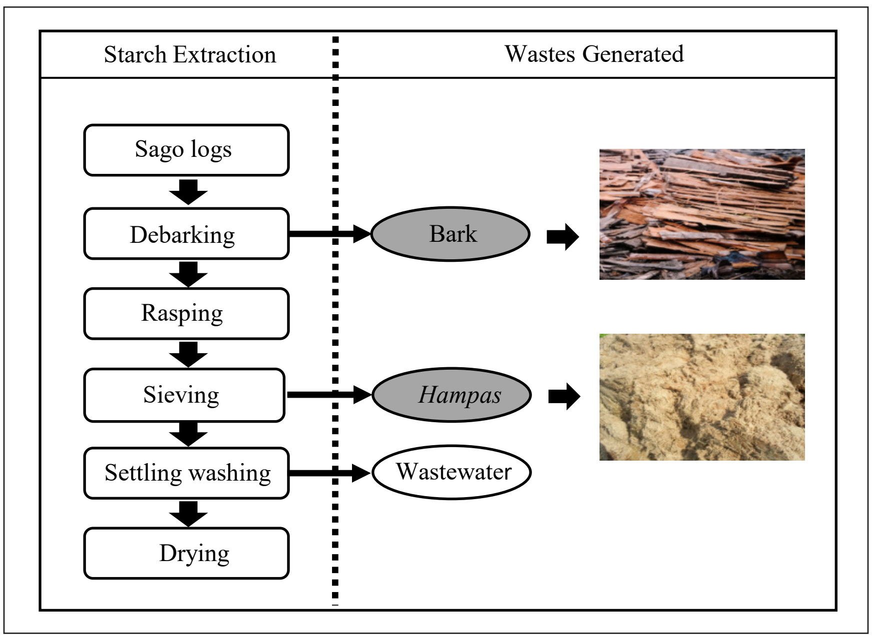


Figure 1. Schematic flow diagram of sago starch extraction

Starch Production Statistics

The Malaysian sago palm industry is one of the crucial sago exporters in the world, exporting more than 40,000 tons of sago starch since 2004. It was reported that the production and export value of sago starch has been increasing 15.0% to 20.0% every year (Jenol et al., 2014; Mohamad et al., 2011). Figure 2 illustrates the 10 years sago starch export trend from 2008 to 2017, as reported by DAS. The sago starch export was at peak, with the export value of more than 47,000 tons from 2011 to 2013, before it dropped to approximately 40,000 tons in four subsequent years, 2014 to 2017. Nonetheless, the export value in 2017 rose a bit which shows the tendency of sago starch export to increase in the future. The displayed trend indicates the global demand for sago starch. When the demand increases, so will sago production, which consequently contributes to the significant amount of waste generated. Hence, the sago industry is facing waste disposal and management problems.

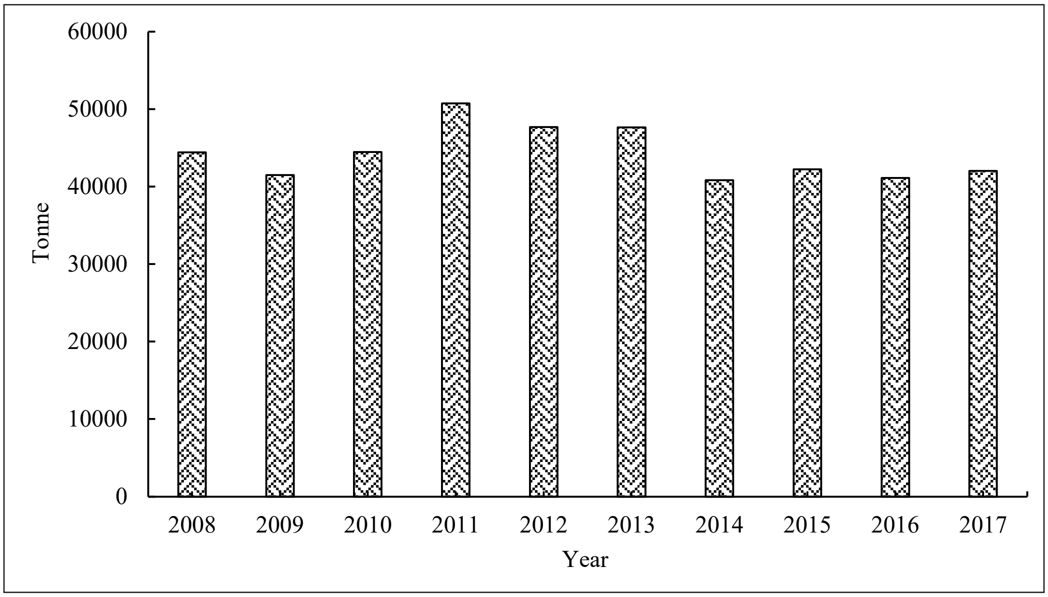


Figure 2. Ten (10) years Malaysia sago starch export statistics, 2008 - 2017

SAGO WASTES

During the processing of sago starch, three major by-products are generated, namely bark of sago trunk, fibrous pith residue, which is also known as *hampas*, and wastewater (Figure 1). Bark and *hampas* are classified as solid residues whereas wastewater is a liquid residue (Awg-Adeni et al., 2010). Sago wastes consist of non-starch polysaccharides (NSP) or lignocellulosic materials, i.e. cellulose, hemicellulose, and lignin. In a day, 15.6 tons of woody bark, 238 tons of wastewater, and 7.10 tons of fibrous pith residue are generated from approximately 600 logs of sago palm (Ngaini et al., 2014a).

Solid Residue

Sago *Hampas.* Sago *hampas* is an inexpensive fibrous residue left behind after the starch extraction process, particularly from crushing and sieving (Figure 1). It is also known as sago pith residue, a starchy lignocellulosic by-product. In Sarawak, mainly in Sibu and Mukah divisions, about 50.0 to 100 tons of sago *hampas* are produced daily (Awg-Adeni et al., 2010). Previous studies have revealed that a large number of starch granules, approximately 60.0%–70.0% starch per gram dry weight of *hampas*, are still trapped within the lignocellulosic matrix (Singhal et al., 2008; Vikineswary et al., 1994). Since a significant amount of starch is left within the *hampas*, it has been used as feedstuffs for swine around sago processing areas.

Hampas is a refuse from sago starch extraction process and is usually discarded into nearby streams together with the factory wastewater (Vincent et al., 2015a). Eventually, the streams' oxygen content, which supports higher life forms, decreases due to the high consumption of dissolved oxygen for degradation of the wastes via microbial activity. Indubitably, this will negatively affect the environment, whereby a severe drop in water quality can endanger aquatic lives (Lai et al., 2013). Due to the presence of lignocellulosic fibrous materials (Table 1), *hampas* has been used as animal feed, compost for mushroom culture, substrate in hydrolysis process to produce confectioners syrup, and for particleboard manufacture (Phang et al., 2000). Apart from that, sago *hampas* is also applied as a cheap carbon source for production of several enzymes via fermentation like laccase (Kumaran et al., 1997), α-amylase, and cellulase (Khan & Husaini, 2006).

Table 1
Chemical composition of sago hampas and SB

Item	**Sago *Hampas***			**SB**		
Reference	Ozawa et al. (1998)	Lee et al. (2014)	Jenol et al. (2014)	Mohamad et al. (2011)	Ethaib et al. (2016)	Erabee et al. (2017)
Starch (%)	49.7	37.0	49.5	-	-	-
Cellulose (%)	20.2	40.0	26.0	44.1	40.8	44.1
Hemicellulose (%)	7.9	13.0	14.5	21.1	22.3	21.1
Lignin (%)	3.3	5.0	7.5	23.3	25.9	23.3

Sago Bark. Basically, there are two types of sago bark (SB) wastes, which are peeled and shredded. Most of the peeled types come from the lower part of a sago tree and are manually peeled using a parang due to its uneven shape. It has a thicker peripheral, approximately 2.00 to 3.00 cm thick. On the other hand, shredded types are produced during the debarking process using machine (Aziin & Rahman, 2005). The bark accounts for about 17.0% of the processed logs and approximately 5.00–15.0 tons of bark per day is produced from a factory (Vikineswary et al., 1994). Annually, more than 20,000 tons of SB are discarded from Malaysia's sago industry.

A lot of SB wastes left behind during the starch extraction process are destroyed through open or controlled burning. This practice poses several serious problems such as threat to the environment through air pollution and furnace chimney corrosion. Discarding SB into rivers is also inappropriate as it could cause the rivers to become shallow. Likewise, leaving tones of SB to naturally degrade would consume time and space besides creating an unpleasant odor and view. Furthermore, only less than 15.0% of SB is used as fuel in sago processing mills due to the corrosion problem caused by SB since the compound is acidic in nature (Wahi et al., 2014). Nevertheless, due to its woody nature, locals usually use SB wastes to build a platform around the factory, footpaths of houses, wall materials, and fences. Moreover, numerous interior decoration products, such as wall tiles, furniture, and flower pots or containers, have been successfully recycled from SB, ascribing to the natural features and beautiful surface of SB (Aziin & Rahman, 2005). Not only that, an eco-friendly material, glulam, from SB is used in combination with less used wood species, *pulai* and *binuang*, in an effort to maximize SB utilization (Wahyudi & Arifudin, 2017). The detailed chemical composition of SB is tabulated in Table 1.

APPLICATION OF SAGO WASTES

Due to the high amount of residual wastes generated every year from the sago industry, efforts have been made to resolve this plaguing issue by valorizing the wastes into value-added materials. The wastes have been widely used as feedstocks or raw materials in industrial applications owing to the large quantity of carbon neutral still embedded within its complex lignocellulosic matrix. Hence, use of sago wastes is one of the solutions to the impending petroleum shortage crisis and a way to address environmental issues by making full use of sustainable resources. The following sections discuss the applications pertinent to the utilization of sago wastes.

Adsorbents

Heavy Metals Adsorbent. Low-cost adsorbents made from agricultural wastes are now gaining more interest than conventional adsorbents to remove harmful contaminants from watercourses because it is abundantly available and economically attractive. Sago *hampas* is largely composed of cellulose and lignin and thus has potential to be used directly as a biosorbent (Vikineswary et al., 1994). Ground, dried, and sieved untreated sago *hampas* has been used to remove metal ion pollutants like copper, Cu(II), and lead Pb(II) in aqueous solution (Quek et al., 1998). The findings showed that sago *hampas* adsorbed Pb(II) better than Cu(II), with a higher initial sorption rate and greater sorption capacity. On the other hand, chemically activated sago *hampas* using H_2SO_4 and $(NH_4)S_2O_8$ efficiently removed both Cu(II) (Maheswari et al., 2008) and Pb(II) (Karthika et al., 2010). In general, a lower pH is more favorable for effective adsorptions of Pb(II) and Cu(II). Nonetheless, a higher pH was found optimum for Zn adsorption by sago waste compared to pH 7 (Wahi et al., 2010). Besides, it was proven that sago *hampas* had the potential to be used as a low-cost biosorbent for mercury removal (Saman et al., 2014). A better adsorption was found for inorganic mercury (Hg(II)) than that of organic mercury ($CH_3Hg(I)$), with a maximum adsorption capacity of 0.288 and 0.213 mmol/g, respectively. Adsorption of mercury ions was highly dependent on the solution's pH in which Hg(II) adsorption was higher at pH > 4, while $CH_3Hg(I)$ adsorption was higher at pH < 5.

Activated Carbon (AC). Various agricultural wastes, including sago *hampas*, have been used for activated carbon (AC) preparation using concentrated sulfuric acid (Kadirvelu, 2003). Sago *hampas* AC significantly adsorbed different dye and metal ion components from aqueous solutions as early as 1 h and registered a more than 90.0% removal in 24 h. Specifically, it effectively adsorbed mercury (Hg(II)) (Kadirvelu et al., 2004) and Rhodamine-B (Kadirvelu et al., 2005) with a 55.6 mg/g and 16.1 mg/g adsorption capacity, respectively, for a particle size range of 125–250 µm at approximately pH 5. Meanwhile, physiochemical activation of SB via $ZnCl_2$ produced a higher surface area AC compared to commercial AC made from coconut shells (Erabee et al., 2017). Low ash and high carbon and oxygen content in waste feedstocks positively affect the production of a good adsorbent

Oil Adsorbent. Oil pollution negatively impacts both the economy and environment and remains a serious concern in Malaysia. Natural organic sorbents, particularly those made from agricultural wastes, have many advantages like having a comparable density with synthetic sorbents, chemical-free, and highly biodegradable (Annunciado et al., 2005). Some are also able to significantly absorb more oil than commercial sorbents made of polypropylene (Adebajo & Frost, 2004). However, most of the cellulose portions are covered by the hydroxyl groups, causing the cellulose to behave more hydrophilic than hydrophobic (Said et al., 2009), hence, will adsorb water and reduce the oil sorption capacities (Ali et al., 2012). In response, numerous studies have been conducted to improve the hydrophobicity and oleophilicity by means of alkalization, salt treatment, surfactant treatment, and esterification (Wahi et al., 2013).

Esterification of SB (Ngaini et al., 2014a) and sago *hampas* (Ngaini, et al., 2014b) using stearic acid (SA) resulted in a higher used engine oil (UEO) sorption capacity and lesser water uptake in an aqueous environment compared to untreated ones. Besides that, the findings also highlighted a higher buoyancy (able to remain afloat for 7 days) and sorption capacity in esterified sago bark (ESB) than other natural sorbents. To enhance ESB's adsorption ability, various parameter relationships were examined (Wahi et al., 2014). ESB with a preparation condition of 1:1 SB:SA, 15.0% catalyst (CaO), and 8 h refluxing time produced a maximum oil removal efficiency of up to 95.5% compared to untreated SB (45.4%). Later, this optimized ESB's capability to remove oil, particularly from palm oil mill effluent (POME), was tested (Wahi et al., 2017). After a systematic investigation of various parameters, ESB was revealed to be a potentially better palm oil sorbent with a consistent 88.0%–89.0% oil removal efficiency and had excellent reusability of up to six cycles. Interestingly, ESB has also demonstrated its capability in removing oil spilled in a dynamic seawater environment (Ngaini et al., 2017). The oil sorption capacity after 60 min was 5.70 g/g in the dynamic seawater system compared to 3.30 g/g in deionized water. This suggests that ESB is a promising low-cost natural sorbent that works best in dynamic seawater conditions.

Sugar Sources

Sago waste can also be utilized to produce fermentable sugars. Sago *hampas* was analyzed to determine its carbohydrate content using cold and hot water extraction and acid hydrolysate (Akmar & Kennedy, 2001). All sago *hampas* hydrolysates showed a significantly higher total monosaccharides content, especially in glucose, compared to oil palm trunk wastes. This makes sago *hampas* a potential fermentation substrate candidate. Optimum glucose production from sago *hampas* through acid and enzymatic hydrolyses were also systematically investigated using H_2SO_4 and glucoamylase, respectively (Kumoro et al., 2008). A maximum of 0.565 and 0.623 g glucose/g waste were generated via enzymatic and acid hydrolyses, respectively. Likewise, hydrolysis of sago *hampas* using microbial enzymes containing 24.0 IU β-glucosidase produced by *Aspergillus fumigatus* UPM2, gave a higher reducing sugar production of 20.8 g/L (Linggang et al., 2012). Meanwhile, a local fungus, *Trichoderma* sp. KUOM0001 was found suitable to be applied on sago *hampas* for reducing sugar conversion via solid substrate fermentation (SSF) (Shahrim et al., 2008). The optimized condition during SSF after 96 h incubation further maximized the generation of reducing sugar i.e. from 24.0 mg/L obtained without optimization to 46.0 mg/L.

In addition, several pre-treatment methods for sago *hampas* were developed for bio-convertible polysaccharides recovery prior to employing enzymatic hydrolysis (Janggu & Bujang, 2009). Steaming treatment produced a significant starch concentration in liquid fraction with a relatively high solid fraction recovery of 50.6%, high convertibility degree of cellulose fiber at 42.0%, and high lignin reduction. This makes it the best pre-treatment compared to the other two alkaline pre-treatments examined in the study. Furthermore, another study demonstrated that ionic liquids could have dual-functions, namely, biomass dissolution and structural disruption, which facilitated maximum recovery of reducing sugars from sago *hampas* (Lee et al., 2014). A maximum reducing sugars recovery of 90.0% was achieved via incorporation of the ionic liquid pre-treatment, solid acid, and enzymatic saccharification, using 1-butyl-3-methylimidazolium chloride ([BMIM]Cl), Amberlyst 15 (A15), and *Trichoderma viride* cellulase, respectively.

Recently, several studies were conducted on SB to test its potential to be used as a raw material for production of sugars. Due to the presence of complex lignocellulosic components, microwave-assisted dilute acid pre-treatment was introduced prior to commencing enzymatic hydrolysis (Ethaib et al., 2016). After the pre-treatment, there was a 32.0% reduction in lignin content, an increased crystallinity from 29.0% to 47.0%, and clear damage and fragmentation to the surface structure of SB. The enzymatic hydrolysis showed that 24.0 FPU/g cellulase and 50.0 U/g β-glucosidase were sufficient to obtain a higher sugar yield from pre-treated SB. Apart from glucose, monosaccharide xylose was also produced from SB. Xylose is a carbon source for microorganism growth to produce xylitol. The xylose concentration achieved was 22.8 g/L under optimum hydrolysis time and acid concentration of 60 min and 8%, respectively (Mohamad et al., 2011).

Biofuels

Bioethanol. The current dependency on starchy biomass to produce ethanol has raised concerns regarding food shortages, fuel conflicts, and serious destruction of soil resources (Balat & Balat, 2009; Kennes et al., 2015). Hence, extensive research efforts have focused on the potential of using inedible feedstocks as alternatives to starchy biomass (Sun & Cheng, 2002). In this regard, lignocellulosic biomass, which does not compete with food production or animal feed, was determined to be an appropriate substrate for bioethanol production (Limayem & Ricke, 2012). Sago *hampas*, which is mainly composed of lignocellulosic materials, has been subjected to three cycles of enzymatic hydrolysis, whereby the hydrolysate is reused, prior to bioethanol production (Awg-Adeni et al., 2013). This method has shown great improvement in sugar concentration and produced a higher ethanol yield of 93.3% per total glucose existing in the fermentation media.

Simultaneous co-saccharification and fermentation (Sc-SF) was also carried out on sago *hampas* using two different enzymes, amylase and cellulase (Vincent et al., 2015b). During the Sc-SF process, both enzymes were added together into the culture for starch and fiber co-saccharification, while *Saccharomyces cerevisiae* was used to ferment the sugars in hydrolysate. A maximum of 37.9 g/L glucose and 80.5% theoretical ethanol yield were obtained from 5.00% sago *hampas* loading. On the other hand, another study introduced a modified version of sago *hampas* Sc-SF called sequential saccharification and simultaneous fermentation (SSSF) (Vincent et al., 2015a). Distinct from Sc-SF, this technique was performed in two steps, i.e. Stage 1 (amylolytic step) and Stage 2 (cellulolytic step). Cellulase was added into the culture sequentially after amylase during the designated 5 days experiment. Although the highest ethanol yield from SSSF of 77.4% was lower than that of Sc-SF, both techniques converged the idea of producing a high ethanol yield when different enzyme mixtures and *S. cerevisiae* were combined in a culture.

Besides that, a recent study had demonstrated the effect of acid hydrolysis on sago *hampas* and compared it with conventional enzymatic hydrolysis (Bukhari et al., 2017). Dilute acid hydrolysis resulted in a substantial amount of fermentable sugars with 96.0% hydrolysis conversion compared to 79.0% achieved using enzymatic approach. Furthermore, the highest yield of 98.0% ethanol was also attained from acid hydrolysate without supplementation of nitrogen and nutrients. Considering the environmental aspect, time consumption, and cost issues posed by both enzymatic and acid hydrolysis methods, production of ethanol has switched towards utilizing emerging advanced technology. For instance, microwave technology has been applied in bioethanol production from sago *hampas* (Thangavelu et al., 2014) and SB (Kannan et al., 2013). These studies have reported promising ethanol yield and limited by-products.

Biobutanol/Bio-Oils/Biohydrogen/Co-Firing Material. Besides bioethanol, biotechnological production of biobutanol has also gained increased interest for use as alternative biofuel. Fermentable sugars from enzymatic hydrolysis of sago *hampas* were converted to acetone-butanol-ethanol (ABE) by *Clostridium acetobutylicum* ATCC 824 (Linggang et al., 2013). This process produced approximately 8.84 ± 0.200 g/L ABE (5.41 ± 0.100 g/L butanol). Bio-oils can also be produced from sago *hampas* via microwave pyrolysis (Mona et al., 2013). A high colorific value of 22.0 MJ/kg was recorded and a significant hydrocarbon quantity was identified from sago *hampas*, enabling it to be used as a liquid fuel. In addition, a higher biohydrogen yield (2.65 mol H_2/mol glucose) was achieved when sago *hampas* was fermented with *Clostridium butyricum* A1, compared to using synthetic glucose as the substrate (1.90 mol H_2/mol glucose) (Jenol et al., 2014). Furthermore, SB has been utilized as a co-firing fuel along with commonly used biomass wastes (Chong et al., 2014). SB recorded a competitive heating value (HV) of 19.0 MJ/kg in the dry state and hence is suitable for co-firing during coal power generation. In a techno-economic study, SB was found to be feasible for use as feedstocks for a combined heat and power (CHP) system. A maximum of 472 kW of net electricity was generated with a payback period of 3.51 years and carbon saving of 5475 kg CO_2/d (Kin et al., 2015).

Nanomaterials

Along with the rising global awareness of creating sustainable communities via waste minimization, sago wastes were found to be reliable low-cost precursors in preparing silica, a material that possesses a strong luminescent property. Nanocrystalline silica xerogel (NSX) powder was produced from an amorphous silica xerogel (ASX) extracted from sago *hampas* ash. The crystalline silica xerogel (CSX) was derived from ASX via calcination process using an electric furnace with a heating rate of 2°C/min and retention time of 8 h at 1,200°C (Aripin et al., 2013). It was discovered that NSX transformed into a stable form at 1,200°C. Likewise, sago waste was also successfully converted into highly fluorescing carbon dots (C-dots) via a simple thermal hydrolysis method without any surface passivation (Tan et al., 2014). The C-dots can be dispersed in aqueous media and it will portray a significant fluorescent property in the presence of various metal ions that can be observed by the naked eye under a UV light source. Moreover, a better quenching effect was observed on both Cu(II) and Pb(II).

Apart from that, SB has been utilized as a carbon precursor in the synthesis of catalyst-free carbon nanospheres (CNs) using a simple one-step pyrolysis technique (Hegde et al., 2015). Synthesized CNs displayed a porous nature and it was revealed that more than 95.0% of the carbon particles were 40.0–70.0 nm in size. The specific capacitance value of 180 Fg^{-1} at 2.00 mVs^{-1} and the cycling stability of up to 1,700 cycles have made the obtained CNs suitable for use in supercapacitor applications. Besides its high supercapacitance value, a continuity study had demonstrated that SB CNs significantly enhanced the tensile strength of basalt fiber via homogenous distribution in the epoxy matrix (Mengal et al., 2016). The results showed a significant improvement in tensile strength when 0.600 wt% and 1.00 wt% of CNs particles were included in the basalt fiber-reinforced epoxy composite laminate. Nevertheless, the particle loading of 1.00 wt% showed the best result when 80.6% tensile strength increment was recorded (679 MPa) in comparison to neat basalt fiber-reinforced epoxy composite laminate (376 MPa).

Ceramics

The abundance and availability of agricultural by-products with high silica content and low levels of inorganic compounds have made them a promising raw material for synthesizing silica xerogel (SX). This was confirmed when CaO and Si_2O were the major compounds found in sago pith waste ash (SPWA) and increased in value with the rise of calcination temperature (Rashid et al., 2016). These by-products are usually inexpensive and their effective utilization is highly desirable from an environmental point of view. SX can be produced using either synthetic or natural materials. Numerous studies have demonstrated that a great number of natural resources were effectively utilized to produce SX (Kalapathy et al., 2000; Okoronkwo et al., 2013; Sudiana et al., 2015). Silica-glass ceramic is produced from SX through the heat treatment process known as sintering (El Hamzaoui et al., 2010; Rao et al., 1999). Intensive studies have been done on SX extracted from sago *hampas* ash used for manufacturing ceramics.

Silica-glass ceramic was successfully produced, using SX extracted from sago *hampas* ash, via microwave thermal treatment using a sub-millimeter wave gyrotron (300 GHz) as the radiation source (Aripin et al., 2011a). SX was found to crystallize in the cristobalite phase at 800°C sintering temperature, which is about 200°C lower than the one observed in the conventional process. This ceramic material reached a bulk density at 900°C (about 2.20 g/cm^3), which is probably the same as that of vitreous silica. In a structural and microwave properties study on SX glass-ceramic subjected to different radiation frequencies (8.20 to 12.4 GHz), the reflection losses decreased as the sintering temperature increased. This was due to the formation of fully crystallized silica-glass ceramic under sufficiently high temperature as it became more transparent to the microwaves (Aripin et al., 2012). Hence, SX's high transparency and good chemical durability to irradiation exposure have made it a good candidate for non-linear optics applications such as optical waveguides (Kim et al., 2005; Xu et al., 2010).

Nonetheless, the applicability of SX is limited by its brittleness and certain structural reliability issues, for instance, poor chemical stability at high temperatures in the presence of water, which results in the formation of surface hydroxyls. Thus, incorporating other properly selected oxides into SX helps to improve its properties. For example, incorporation of TiO_2 into SX gives an appreciable effect on the crystallite size of TiO_2 and formation of Si-O-Ti bonds at 900°C (Aripin et al., 2016a). On the other hand, inclusion of SnO_2 increased the bulk density of silica-based glass ceramics and its crystallization, while simultaneously reducing its porosity (Aripin et al., 2016b). Different compositions of Al_2O_3 mixed with SX substantially influenced mullite crystallization of ceramics (Aripin et al., 2013). Results showed that the intensities of cristobalite peaks gradually reduced, while those of the characteristics peaks of mullite amplified with increasing Al_2O_3 content. In addition, another study had demonstrated that the mixture of sago waste ash with clay had successfully produced highly porous ceramics (Aripin et al., 2011b). The results showed 10.0 wt% ash could be incorporated into the clay material without any adverse effects on the bulk density and porosity of the samples sintered at 1,000 °C. Clay ceramic with a high porosity can be applied for particulates removal from hot gas stream and catalyst recovery from off-gases and, hence, is suitable for industrial application.

Composites

Most biodegradable plastics rely on cheap and abundant by-products from the agro-industry, which have a faster biodegradable rate and are relatively hydrophilic in nature and somewhat crystalline in structure. Based on preceding works, renewable biodegradable polymer used alone tends to cause performance and processing problems; hence, blending it with synthetic polymers such as polyvinyl alcohol (PVA) is highly desirable. PVA is the world's largest synthetic water-soluble and biodegradable polymer (Ramaraj, 2007) with an outstanding chemical resistance, tensile strength, flexibility, and barrier properties. Due to its high price, PVA has been restricted from being substantially used in applications (Schellekens & Bastiaansen, 1991). Nonetheless, in a different compounding methods study, incorporation of sago *hampas* into PVA yielded the highest value of tensile strength and elongation at break, particularly for dry-blending PVA with sago *hampas* (PVA/SPW) and pre-plasticized PVA and sago *hampas* (pPVA/pSPW) (Toh et al. 2011a). In addition, another study (Toh et al., 2011b) showed that blending sago *hampas* with a high PVA loading created great molecular force interactions that could overtake those of PVA itself. This outcome indicates that the interaction between PVA and sago *hampas* is reactive, thus inducing synergistic effects.

In accordance with the high starch content in sago *hampas*, plasticization of this waste is believed to produce biodegradable composite materials without the addition of any synthetic plastics as a binder. This characteristic makes the sago *hampas* a unique and special material compared to other fibrous wastes. With glycerol and water as plasticizers, sago *hampas* waste was successfully plasticized to form a natural fiber filled thermoplastic using a twin screw extruder (Lai et al., 2013). Plasticization causes the disruption of the original C-type crystallinity. However, V_H and B types of crystallinities were developed due to the reorganization of amylose, amylose-lipid complex (very fast), and amylopectin chains (slow). Nevertheless, increasing glycerol content adversely affected the tensile strength of the plasticized sago *hampas*, while elongation at break remained unchanged (Lai et al., 2014). Phase separation and fiber agglomeration were determined to be responsible contributors to this trend. Distinct molecular restructuring took place in plasticized sago *hampas* upon ageing, which was reflected in the bulk properties leading to an increase in the tensile strength and a decline in elongation at break (Lai et al., 2016). Glycerol and fiber were the two compounds that restrained the retrogradation process of plasticized sago *hampas*. Besides that, a study was also conducted to investigate the potential of SB as a reinforcement material for a natural fiber composite (Abral et al., 2012). Treated SB using 5.00% NaOH after mercerization revealed not only an elastic area but also a plastic zone. The strength of treated bark fiber significantly increased whereby the average ultimate strength was 163 MPa compared to 46.0 MPa of the untreated bark fiber. Similarly, a greater elastic modulus was recorded for treated bark fiber than untreated bark fiber.

A Summary of Sago *Hampas* and SB Applications

Table 2 summarizes and differentiates all the discussed sago wastes applications. Several studies on solid sago wastes have been done recently, indicating the increase in awareness to fully utilize them as value-added materials. Nevertheless, more rigorous researches were conducted on sago *hampas* than SB. Such a difference emphasizes the lower amount of complex lignocellulosic materials and higher starch content in sago *hampas* compared to SB, which had likely drawn much of the researchers' attention. Despite that, there are many studies (Table 2) that have successfully utilized SB with outcomes that are comparable to sago *hampas*.

Table 2
Summary of recent sago wastes applications

Sago waste	Application	Source
Sago *hampas*	**1. Adsorbents**	
	Heavy metals	Karthika et al. (2010); Maheswari et al. (2008); Quek et al. (1998); Saman et al. (2014); Wahi et al. (2010)
	Activated carbon (AC)	Kadirvelu (2003); Kadirvelu et al. (2005, 2004)
	Oil	Ngaini et al. (2014b)
	2. Sugar sources	Akmar and Kennedy (2001); Janggu and Bujang (2009); Kumoro et al. (2008); Lee et al. (2014); Linggang et al. (2012); Shahrim et al. (2008)

Table 2 *(continue)*

Sago waste	Application	Source
Sago *hampas*	**3. Biofuels**	
	Bioethanol	Awg-Adeni et al. (2013); Bukhari et al. (2017); Thangavelu et al. (2014); Vincent et al. (2015b); Vincent et al. (2015a)
	Biobutanol	Linggang et al. (2013)
	Bio-oil	Mona et al. (2013)
	Biohydrogen	Jenol et al. (2014)
	4. Nanomaterials	Aripin et al. (2013); Tan et al. (2014)
	5. Ceramics	Aripin et al. (2011b); Aripin et al. (2013, 2012); Aripin et al. (2011a); Aripin et al. (2016a); Aripin et al. (2016b); Rashid et al. (2016)
	6. Composites	Lai et al. (2016, 2013, 2014); Toh et al. (2011a); Toh et al. (2011b)
Sago bark	**1. Adsorbents**	
	Activated carbon (AC)	Erabee et al. (2017)
	Oil	Ngaini et al. (2017); Ngaini et al. (2014a); Wahi et al. (2017, 2014)
	2. Sugar sources	Ethaib et al. (2016); Mohamad et al. (2011)
	3. Biofuels	
	Bioethanol	Kannan et al. (2013)
	Co-firing	Chong et al. (2014); Kin et al. (2015)
	4. Nanomaterials	Hegde et al. (2015); Mengal et al. (2016)
	5. Composites	Abral et al. (2012)

CONCLUSIONS

Sago wastes are receiving more attention because these residual solids are a viable and utilizable resource for conversion into value-added products. Sago *hampas* and bark have high potential to be used as sustainable feedstocks due to its availability and significant content of starchy-lignocellulosic compounds. When chemically-treated, the sago wastes were able to be a good adsorbent for both heavy metals and oils. Enzymatic and acid hydrolysis approached capable of extracting sugars from the polysaccharides matrix of the sago wastes. Also, the fermentable sugars produced i.e. glucose were used to generate bioethanol, bio-butanol and bio-hydrogen via fermentation using different bacteria accordingly. Meanwhile, high temperature thermal treatment enhanced the special nature of sago wastes, hence producing good nanomaterials products i.e. C-dots and CNs. The high silica content and low level of inorganic compounds contained in sago wastes have made it as a promising source in producing silica-glass ceramic. Therefore, it can be deduced that conversion of sago wastes into useful products is a promising approach to mitigate environmental pollution while concurrently reducing the dependency on petroleum resources. However, among the solid sago wastes generated, emphasis should also be given to SB, parallel with the rigorous studies done on sago *hampas,* due to its high cellulose and hemicellulose content, regardless of the presence of lignin. Newer chemical syntheses and technological adaptations should be considered to overcome the strong lignin structure that is strongly bound with precious neutral carbon, i.e. hemicellulosic materials.

ACKNOWLEDGEMENT

Authors are grateful for the fund received from Putra Grant – Putra Graduate Initiative (GP-IPS 9536400) from Universiti Putra Malaysia.

REFERENCES

Abd-Aziz, S. (2002). Sago starch and its utilisation. *Journal of Bioscience and Bioengineering*, *94*(6), 526-529.

Aziin, K. A., & Rahman, A. (2005, December 12-14). Utilizing sago (metroxylon spp) bark waste for value added products. In *2005 4th International Symposium on Environmentally Conscious Design and Inverse Manufacturing* (pp. 102-106). Tokyo, Japan.

Abral, H., Kasmianto, E., & Perdana, M. (2012). Mechanical properties and microstructure of metroxylon sago fiber treated by sodium hydroxide. *International Journal of Technology*, *1*, 16-23.

Adebajo, M. O., & Frost, R. L. (2004). Acetylation of raw cotton for oil spill cleanup application: An FTIR and 13C MAS NMR spectroscopic investigation. *Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy*, *60*(10), 2315-2321.

Akmar, P. F., & Kennedy, J. F. (2001). The potential of oil and sago palm trunk wastes as carbohydrate resources. *Wood Science and Technology*, *35*(5), 467-473.

Ali, N., El-Harbawi, M., Jabal, A. A., & Yin, C. Y. (2012). Characteristics and oil sorption effectiveness of kapok fibre, sugarcane bagasse and rice husks: Oil removal suitability matrix. *Environmental Technology*, *33*(4), 481-486.

Annunciado, T. R., Sydenstricker, T. H. D., & Amico, S. C. (2005). Experimental investigation of various vegetable fibers as sorbent materials for oil spills. *Marine Pollution Bulletin*, *50*(11), 1340-1346.

Aripin, H., Mitsudo, S., & Sudiana, I. N. (2011a). Rapid sintering of silica xerogel ceramic derived from sago waste ash using sub-millimeter wave heating with a 300 GHz CW gyrotron. *Journal of Infrared, Millimeter, and Terahertz Waves*, *32*(6), 867-876.

Aripin, H., Lestari, L., Agusu, L., Sudiana, I. N., & Jumsiah, N. (2011b). Preparation of porous ceramic with controllable additive and firing temperature. *Advanced Materials Research*, *277*, 151-158.

Aripin, H., Mitsudo, S., Endang S, P., Sudiana, I. N., Kikuchi, H., Sano, S., & Sabchevski, S. (2013). Structural characterization of mullite-based ceramic material from Al_2O_3 and silica xerogel converted from sago waste ash. *Advanced Materials Research*, *789*, 262-268.

Aripin, H., Mitsudo, S., Prima, E. S., Sudiana, I. N., Tani, S., Sako, K., & Fujii, Y. (2012). Structural and microwave properties of silica xerogel glass-ceramic sintered by sub-millimeter wave heating using a gyrotron. *Journal of Infrared, Millimeter, and Terahertz Waves*, *33*(11), 1149-1162.

Aripin, H., Mitsudo, S., Sudiana, I. N., Busaeri, N., Sunendar, B., & Sabchevski, S. (2016a). Structural characterization of a glass ceramic developed from TiO_2 and a novel material-silica xerogel converted from sago waste ash. *Materials Science Forum*, *872*, 81-86.

Aripin, H., Mitsudo, S., Sudiana, I. N., Priatna, E., Kikuchi, H., & Sabchevski, S. (2016b). Densification behavior of SnO_2-glass composites developed from the incorporation of silica xerogel and SnO_2. *International Journal of Technology*, *3*, 401-407.

Aripin, H., Mitsudo, S., Sudiana, I. N., Prima, E. S., & Sako, K. (2013). Microstructural and thermal properties of nanocrystalline silica xerogel powders converted from sago waste ash material. *Materials Science Forum*, *737*, 110-118.

Awg-Adeni, D. S., Abd-Aziz, S., Bujang, K., & Hassan, M. A. (2010). Bioconversion of sago residue into value added products. *African Journal of Biotechnology*, *9*(14), 2016-2021.

Awg-Adeni, D. S., Bujang, K. B., Hassan, M. A., & Abd-Aziz, S. (2013). Recovery of glucose from residual starch of sago *hampas* for bioethanol production. *BioMed Research International*, *2013*, 1-8.

Balat, M., & Balat, H. (2009). Recent trends in global production and utilization of bio-ethanol fuel. *Applied Energy*, *86*(11), 2273-2282.

Bujang, K. (2006). Potentials of bioenergy from the sago industries in Malaysia. In H. W. Doelle, S. Rokem & M. Berovic (Eds.), *EOLSS - Encyclopedia of Life Support Systems* (pp. 124-131). Oxford, UK: EOLSS Publisher.

Bukhari, N. A., Loh, S. K., Nasrin, A. B., & Maizan, I. (2017). Hydrolysis of residual starch from sago pith residue and its fermentation to bioethanol. *Sains Malaysiana*, *46*(8), 1269-1278.

Chew, T. A., Isa, A. H. M., & Mohayidin, M. G. (1999). Sago (Metroxylon sagu Rottboll), the forgotten palm. *Journal of Sustainable Agriculture*, *14*(4), 5-17.

Chew, T., Hassan, A., Ghazah, M., & Ghazah, I. (1998, October 14-19). The sago industry in Malaysia : Present status and future prospects. In *Proceedings of the 7th international Working Conference on Stores-product protection* (Vol. 2, pp. 1720-1728). Beijing, China.

Chong, K. H., Law, P. L., Rigit, A. R. H., Baini, R., & Shanti, F. S. (2014). Sago bark as renewable energy. *UNIMAS E-Journal of Civil Engineering*, *5*(2), 29-34.

Department of Agriculture Sarawak. (2013). *Sarawak Agriculture Statistic*. Retrieved December 20, 2017, from http://www.doa.sarawak.gov.my/modules/web/pages.php?mod=webpage&sub=page&id=712

Erabee, I. K., Ahsan, A., Daud, N. N. N., Idrus, S., Shams, S., Din, M. F. M., & Rezania, S. (2017). Manufacture of low-cost activated carbon using sago palm bark and date pits by physiochemical activation. *BioResources*, *12*(1), 1916-1923.

Ethaib, S., Omar, R., Mazlina, M. K. S., Radiah, A. B. D., & Syafiie, S. (2016). Microwave-assisted dilute acid pretreatment and enzymatic hydrolysis of sago palm bark. *BioResources*, *11*(3), 5687-5702.

El Hamzaoui, H., Courthéoux, L., Nguyen, V. N., Berrier, E., Favre, A., Bigot, L., … & Capoen, B. (2010). From porous silica xerogels to bulk optical glasses : The control of densification. *Materials Chemistry and Physics*, *121*, 83-88.

Flach, M., & Schuiling, D. L. (1988). Revival of an ancient starch crop: A review of the agronomy of the sago palm. *Agroforestry Systems*, *7*(3), 259-281.

Flores, D. M. (2008). The versatile sago (*Metroxylon sagu* Rottb.) and its green potential for Mindanao. *Banwa*, *5*(1), 8-17.

Hegde, G., Manaf, S. A. A., Kumar, A., Gomaa, A. M. A., Chong, K. F., Zainab, N., & Sharma, K. V. (2015). Biowaste sago bark based catalyst free carbon nanospheres : Waste to wealth approach. *ACS Sustainable Chemistry and Engineering*, *3*(9), 2247-2253.

Isikgor, F. H., & Becer, C. R. (2015). Lignocellulosic biomass: A sustainable platform for the production of bio-based chemicals and polymers. *Polymer Chemistry*, *6*(25), 4497-4559.

Janggu, U., & Bujang, K. (2009, April 14-15). Development of pre-treatment of sago fibers for recovery of bio-convertible polysaccharides (cellulose) for ethanol fermentation. In *Proceedings of the 2nd Biotechnology Colloquium 14th-15th April 2009* (pp. 85-90). Universiti Malaysia Sarawak, Malaysia.

Jenol, M. A., Ibrahim, M. F., Yee, P. L., Salleh, M. M., & Abd-Aziz, S. (2014). Sago biomass as a sustainable source for biohydrogen production by *Clostridium butyricum* A1. *BioResources*, *9*(1), 1007-1026.

Kadirvelu, K. (2003). Utilization of various agricultural wastes for activated carbon preparation and application for the removal of dyes and metal ions from aqueous solutions. *Bioresource Technology*, *87*, 129-132.

Kadirvelu, K., Karthika, C., Vennilamani, N., & Pattabhi, S. (2005). Activated carbon from industrial solid waste as an adsorbent for the removal of Rhodamine-B from aqueous solution: Kinetic and equilibrium studies. *Chemosphere*, *60*(8), 1009-1017.

Kadirvelu, K., Kavipriya, M., Karthika, C., Vennilamani, N., & Pattabhi, S. (2004). Mercury (II) adsorption by activated carbon made from sago waste. *Carbon*, *42*(4), 745-752.

Kalapathy, U., Proctor, A., & Shultz, J. (2000). A simple method for production of pure silica from rice hull ash. *Bioresource Technology*, *73*, 257-262.

Kamal, S. M. M., Mahmud, S. N., Hussain, S. A., & Ahmadun, F. R. (2007). Improvement on sago flour processing. *International Journal of Engineering and Technology*, *4*(1), 8-14.

Kannan, S. T., Ahmed, A. ., & Ani, F. (2013). Energy efficient microwave irradiation of sago bark waste (SBW) for bioethanol production. *Advanced Materials Research*, *701*, 249-253.

Karim, A. A., Tie, A. P. L., Manan, D. M. A., & Zaidul, I. S. M. (2008). Starch from the sago (*Metroxylon sagu*) palm tree properties, prospects, and challenges as a new industrial source for food and other uses. *Comprehensive Reviews in Food Science and Food Safety*, *7*(3), 215-228.

Karthika, C., Vannilamani, N., Pattabhi, S., & Sekar, M. (2010). Utilization of sago waste as an adsorbent for the removal of Pb(II) ion from aqueous solution: Kinetic and isotherm studies. *International Journal of Engineering Science and Technology*, *2*(6), 1867-1879.

Kennes, D., Abubackar, N., Diaz, M., Veiga, C., & Kennes, C. (2015). Bioethanol production from biomass : Carbohydrate vs syngas fermentation. *Journal of Chemical Technology and Biotechnology*, *91*, 304-317.

Khan, F. A. B. A., & Husaini, A. A. S. A. (2006). Enhancing α-amylase and cellulase in vivo enzyme expressions on sago pith residue using Bacilllus amyloliquefaciens UMAS 1002. *Biotechnology*, *5*(3), 391-403.

Kim, Y. T., Kim, D. S., & Yoon, D. H. (2005). PECVD SiO_2 and SiON films dependant on the rf bias power for low-loss silica waveguide. *Thin Solid Films*, *475*, 271-274.

Kin, Y., Sadhukhan, J., Siew, K., & Ng, D. K. S. (2015). Techno-economic evaluations for feasibility of sago-based biorefinery , Part 1 : Alternative energy systems. *Chemical Engineering Research and Design*, *107*, 263-279.

Kumaran, S., Sastry, C. A., & Vikineswary, S. (1997). Solid substrate fermentation of sago "hampas" for laccase production. *Studies in Environmental Science*, *66*, 239-248.

Kumoro, A. C., Ngoh, G. C., Hasan, M., Ong, C. H., & Teoh, E. (2008). Conversion of fibrous sago (Metroxylon sagu) waste into fermentable sugar via acid and enzymatic hydrolysis. *Asian Journal of Scientific Research*, *1*(4), 412-420.

Lai, J. C., Rahman, W. A. W. A., Avérous, L., & Lim, T. H. (2016). Study and characterisation of the post processing ageing of sago pith waste biocomposites. *Sains Malaysiana*, *45*(4), 633-641.

Lai, J. C., Rahman, W. A. W. A., & Toh, W. Y. (2013). Characterisation of sago pith waste and its composites. *Industrial Crops and Products*, *45*, 319-326.

Lai, J. C., Rahman, W. A. W. A., & Toh, W. Y. (2014). Mechanical, thermal and water absorption properties of plasticised sago pith waste. *Fibers and Polymers*, *15*(5), 971-978.

Lee, K. M., Ngoh, G. C., Seak, A., Chua, M., Yoon, L. W., Ang, T. N., & Lee, M. (2014). Comparison study of different ionic liquid pretreatments in maximizing total reducing sugars recovery. *BioResources*, *9*(1), 1552-1564.

Limayem, A., & Ricke, S. C. (2012). Lignocellulosic biomass for bioethanol production: Current perspectives, potential issues and future prospects. *Progress in Energy and Combustion Science*, *38*(4), 449-467.

Linggang, S., Phang, L. Y., & Wasoh, H. (2013). Acetone-butanol-ethanol production by Clostridium acetobutylicum ATCC 824 using sago pith residues hydrolysate. *Bioenergy Research*, *6*(1), 321-328.

Linggang, S., Phang, L. Y., Wasoh, M. H., & Abd-Aziz, S. (2012). Sago pith residue as an alternative cheap substrate for fermentable sugars production. *Applied Biochemistry and Biotechnology*, *167*(1), 122-131.

Maheswari, P., Venilamani, N., Madhavakrishnan, S., Shabudeen, P. S. S., Venckatesh, R., & Pattabhi, S. (2008). Utilization of sago waste as an adsorbent for the removal of Cu(II) ion from aqueous solution. *E-Journal of Chemistry*, *5*(2), 233-242.

Mengal, A. N., Akilu, S., & Siddique, M. (2016). Strength enhancement of basalt fiber-reinforced epoxy laminates with biowaste catalyst free carbon nanospheres. *Journal of Applied and Emerging Sciences*, *6*(2), 105-110.

Mohamad, N. L., Kamal, S. M. M., & Abdullah, A. G. L. (2011). Optimization of xylose production from sago trunk cortex by acid hydrolysis. *African Journal of Food Science and Technology*, *2*(5), 102-108.

Mona, S., Aziz, A., Wahi, R., Ngaini, Z., & Hamdan, S. (2013). Bio-oils from microwave pyrolysis of agricultural wastes. *Fuel Processing Technology*, *106*, 744-750.

Naim, H. M., Yaakub, A. N., Hamdan, A., & Asmah, D. (2016). Commercialization of sago through estate plantation scheme in Sarawak: The way forward. *International Journal of Agronomy*, *2016*, 1-6.

Ngaini, Z., Noh, F., & Wahi, R. (2014a). Esterified sago waste for engine oil removal in aqueous environment. *Environmental Technology (United Kingdom)*, *35*(22), 2761-2766.

Ngaini, Z., Wahi, R., Halimatulzahara, D., & Yusoff, N. A. N. (2014b). Chemically modified sago waste for oil adsorption. *Journal of Science and Technology*, *22*(1), 153-161.

Ngaini, Z., Noh, F., & Wahi, R. (2017). Facile sorbent from esterified cellulosic sago waste for engine oil removal in marine environment. *International Journal of Environmental Science and Technology, 15*(2), 341-348.

Okoronkwo, E. A., Imoisili, P. E., & Olusunle, S. O. O. (2013). Extraction and characterization of amorphous silica from corn cob ash by sol-gel method. *Chemistry and Materials Research*, *3*(4), 68-72.

Ozawa, T., Ueno, T., Negishi, O., & Masaki, S. (1998). Chemical characteristics of hemicelluloses in the fibrous residue of sago palm. *Journal of Tropical Agriculture*, *42*(3), 172-178.

Phang, S. M., Miah, M. S., Yeoh, B. G., & Hashim, M. A. (2000). Spirulina cultivation in digested sago starch factory wastewater. *Journal of Applied Phycology*, *12*, 395-400.

Quek, S. Y., Wase, D. A. J., & Forster, C. F. (1998). The use of sago waste for the sorption of lead and copper. *Water SA*, *24*(3), 251-256.

Ramaraj, B. (2007). Crosslinked poly(vinyl alcohol) and starch composite films. II. Physicomechanical, thermal properties and swelling studies. *Journal of Applied Polymer Science*, *103*, 909-916.

Rao, A. V., Wagh, P. B., Haranath, D., Risbud, P. P., & Kumbhare, S. D. (1999). Influence of temperature on the physical properties of TEOS silica xerogels. *Ceramics International*, *25*, 505-509.

Rashid, M., Rohaidzat, M., Johari, M., Azmi, M., & Ahmad, Z. A. (2016). Sago pith waste ash as a new alternative raw materials from agricultural waste. *Materials Science Forum*, *840*, 389-393.

Said, A. E. A. A., Ludwick, A. G., & Aglan, H. A. (2009). Usefulness of raw bagasse for oil absorption: A comparison of raw and acylated bagasse and their components. *Bioresource Technology*, *100*(7), 2219-2222.

Saman, N., Johari, K., Tien, S. S., & Mat, H. (2014). Removal of Hg(II) and $CH_3Hg(I)$ using rasped pith sago residue biosorbent. *Clean - Soil, Air, Water*, *42*(11), 1541-1548.

Schellekens, R., & Bastiaansen, C. (1991). The drawing behavior of polyvinylalcohol fibers. *Journal of Applied Polymer Science*, *43*, 2311-2315.

Shahrim, Z., Sabaratnam, V., Rahman, N. A. A., Abd-Aziz, S., Hassan, M. ., & Karim, M. I. (2008). Production of reducing sugars by Trichoderma sp. KUPM0001 during solid substrate fermentation of sago starch processing waste hampas. *Research Journal of Microbiology*, *3*(9), 569-579.

Singhal, R. S., Kennedy, J. F., Gopalakrishnan, S. M., Kaczmarek, A., Knill, C. J., & Akmar, P. F. (2008). Industrial production, processing, and utilization of sago palm-derived products. *Carbohydrate Polymers*, *72*(1), 1-20.

Sudiana, N., Mitsudo, S., Nishiwaki, T., Susilowati, P. E., Lestari, L., Firihu, M. Z., & Aripin, H. (2015). Effect of microwave radiation on the properties of sintered oxide ceramics. *Contemporary Engineering Sciences*, *8*(34), 1607-1615.

Sun, Y., & Cheng, J. (2002). Hydrolysis of lignocellulosic materials for ethanol production : A review. *Bioresource Technology*, *83*, 1-11.

Tan, X. W., Romainor, A. N., Chin, S. F., & Ng, S. M. (2014). Carbon dots production via pyrolysis of sago waste as potential probe for metal ions sensing. *Journal of Analytical and Applied Pyrolysis*, *105*, 157-165.

Thangavelu, S. K., Ahmed, A. S., & Ani, F. N. (2014). Bioethanol production from sago pith waste using microwave hydrothermal hydrolysis accelerated by carbon dioxide. *Applied Energy*, *128*, 277-283.

Toh, W. Y., Lai, J. C., & Aizan, W. W. (2011a). Influence of compounding methods on poly(vinyl) alcohol/sago pith waste biocomposites: Mechanical and water absorption properties. *Sains Malaysiana*, *40*(7), 719-724.

Toh, W. Y., Sin, L. T., Rahman, W. A. W. A., & Samad, A. A. (2011b). Properties and interactions of poly(vinyl alcohol)-sago pith waste biocomposites. *Journal of Composite Materials*, *45*(21), 2199-2209.

Vikineswary, S., Shim, Y. L., Thambirajah, J. J., & Blakebrough, N. (1994). Possible microbial utilization of sago processing wastes. *Resources, Conservation and Recycling*, *11*(1-4), 289-296.

Vincent, M., Senawi, B. R. A., Esut, E., Nor, N. M., & Awg-Adeni, D. S. (2015a). Sequential saccharification and simultaneous fermentation (SSSF) of sago *hampas* for the production of bioethanol. *Sains Malaysiana*, *44*(6), 899-904.

Vincent, M., Jabang, E., Nur, N. M., Esut, E., Unting, L. B., & Adeni, D. S. A. (2015b). Simultaneous co-saccharification and fermentation of sago *hampas* for bioethanol production. *Agricultural Engineering International: CIGR Journal*, *17*(2), 160-167.

Wahi, R., Abdullah, L. C., Mobarekeh, M. N., Ngaini, Z., & Shean Yaw, T. C. (2017). Utilization of esterified sago bark fibre waste for removal of oil from palm oil mill effluent. *Journal of Environmental Chemical Engineering*, *5*(1), 170-177.

Wahi, R., Chuah, L. A., Choong, T. S. Y., Ngaini, Z., & Nourouzi, M. M. (2013). Oil removal from aqueous state by natural fibrous sorbent: An overview. *Separation and Purification Technology*, *113*, 51-63.

Wahi, R., Chuah, L. A., Ngaini, Z., Nourouzi, M. M., & Choong, T. S. Y. (2014). Esterification of M. sagu bark as an adsorbent for removal of emulsified oil. *Journal of Environmental Chemical Engineering*, *2*(1), 324-331.

Wahi, R., Kanakaraju, D., & Yusuf, N. A. (2010). Preliminary study on zinc removal from aqueous solution by sago wastes. *Global Journal of Environmental Research*, *4*(2), 127-134.

Wahyudi, & Arifudin, M. (2017). Flexural strength properties of glulam made from combining of sago bark and two wood species from West Papua. *European Journal of Wood and Wood Products, 76*(1), 385-388.

Xu, H., Pan, Y., Kou, H., Zhu, Y., & Guo, J. (2010). Preparation of transparent ordered mesoporous carbon/silica composites and their optical limiting properties. *Journal of Alloys and Compounds*, *502*, 6-9.

Pertanika J. Sci. & Technol. 27 (4): 1863 - 1880 (2019)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Optimization of Process parameters for Decolorization of Azo Dye Remazol Golden Yellow by *Bacillus firmus* using Biostatistical Designs

Palanivelan Ramachandran[1]*, Subashni Bhoopathy[2], Ramya Suseenthar[3] and Ayyasamy Pudukkadu Munusamy[1]

[1]*Department of Microbiology, Periyar University, Salem 636 011, Tamilnadu, India*
[2]*Centre for Ocean Research, Sathyabama University, Chennai 600 119, Tamilnadu, India*
[3]*Department of Microbiology, DKM College for Women, Vellore 632 001, Tamilnadu, India*

ABSTRACT

A halo tolerant dye decolorizing bacterium *Bacillus firmus* (TSL9) was isolated from activated textile sludge and identified by 16S rRNA sequencing method. Due to the effect of casein enzymic hydrolysate, and yeast extract, a maximum of 97.23% remazol golden yellow decolorization was manifested by the strain in Luria Bertani medium. Bacterial dye decolorization was insignificant in the aqueous medium, when carbon and nitrogen sources were absent. Plackett-Burman experiments were carried out to screen the significance of factors like beef extract, size of the inoculum and pH on decolorization. Important factors were optimized at various levels by response surface methodology with central composite design. We concluded through model validation that the optimal values of the ingredients required to perform 78.33% decolorization are as follows: beef extract 0.48% (w/v), the size of the inoculum 7.71% (v/v) and pH 6.96. The obtained result specifies that the native bacterium had positive signs to perform effective dye decolorization under derived optimal conditions.

Keywords: 16S rRNA, carbon sources, decolorization, nitrogen sources, optimization, remazol golden yellow

ARTICLE INFO

Article history:
Received: 22 March 2018
Accepted: 17 May 2019
Published: 21 October 2019

E-mail addresses:
starvel2005@yahoo.co.in (Palanivelan Ramachandran)
subashni91@gmail.com (Subashni Bhoopathy)
ramyasuseenthar@gmail.com (Ramya Suseenthar)
pmayyasamy@gmail.com (Ayyasamy Pudukkadu Munusamy)
*Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

INTRODUCTION

Textiles industries are known to be the largest water consumers and one of the imperative sources for the global economy and environmental pollution. Among different industrial sectors, they release a prodigious amount of waste water into the environment (Mondal et al., 2017). In textile industrial practices, the fabric dyeing process contributes to major environmental problems as it involves the utilization of numerous chemicals like lubricants, sequestering agents, dye stuff, soda ash, sodium chloride, acetic acid, soap, fixing and softener (Arslan-Alaton et al., 2008; Julkapli et al., 2014). Dyes are synthetic complex aromatic colored compounds that are more toxic in nature and difficult to degrade, more than 10-90% of dye does not bind with fiber are entered into the environment and sewage treatment plant (Bhattacharya et al., 2018). Consequently investigation on decolorization and degradation of dye molecules from waste water has increased worldwide in the last few decades (Jadhav et al., 2008; Morales-Alvarez et al., 2018).

Among the 12 classes of chromogenic groups of dyes released into the environment, 70% are azo dyes (Tony et al., 2009). Azo dyes, are the major colorants used in textile industries worldwide, and improper disposal of azo dyes cause serious threat to both environment and aquatic organism. The effluent from textile industries containing azo dye causes eutrophication, rapid depletion of dissolved oxygen level in surface water and ground water when released into the environment (Rawat et al., 2016; Solis et al., 2012). Physical and chemical technologies such as photolysis, flocculation, membrane filtration, ultrafiltration, advanced oxidation, electrophotocatalysis, and coagulation with alum, ferric chloride, magnesium, carbon, polymer, mineral sorbents or biosorbents have been studied for the treatment of azo dyes in waste water (Mojsov et al., 2016; Robinson et al., 2001). Although these techniques are effective in removing dyes from contaminated water, they are very expensive for pilot scale operation with a limited potential application leads to the production of the vast quantity of sludge and land pollution. Considering these limitations with regard to chemical and physical processes, the most versatile and widely used technology is biological method (Morales-Alvarez et al., 2018).

Biological decolorization of dyes occur in two ways; one is adsorption on the microbial biomass and another is biodegradation of dyes using bacteria, fungi, actinomycetes and yeasts in aerobic, anaerobic, anaerobic / aerobic (sequential) treatment processes (Bhattacharya et al., 2018; Liu et al., 2006). Biodegradation of dyes is currently viewed as an effective, specific, lower energy-demanding and environmentally benign method. Bacterial decolorization is nonspecific and faster compared to other microorganisms (Kalyani et al., 2009). The biological treatment of textile effluent also has numerous challenges, majorly the capacity of microorganisms to utilize dye as a substrate should be considered (Saratale et al., 2011). Predominantly, azo dyes are chemically stable due to binding of azo group with aromatic amines being difficult to biodegrade aerobically and cannot be utilized as a carbon source by bacteria and it requires additional co-metabolite for the decolorization of dye. Azo dye decolorization is effective in an anaerobic/static condition compared to the aerobic shaking condition, while azoreducatase enzyme has higher activity in reductive cleavage of azo groups (Khehra et al., 2005; Mojsov et al., 2016).

However, the effectiveness of biological treatment system is greatly influenced by various operational parameters such as level of salts, aeration, nutrients, initial concentration of dye, temperature, pH and the amount of inoculum. Operating parameters play a crucial role in the decolorization activity of microorganism (Mondal et al., 2017). Therefore, the effect of various parameters on the color removal process must be investigated and optimized to produce the maximum rate of dye decolorization. The conventional OVAT approach implemented to investigate the media parameters for an ameliorated dye decolorization was unsatisfactory as it was circuitous and needs numerous trials (Bhavsar et al., 2018). Recently bio-statistical tools including design of experiments (DOEs) with response surface methodology (RSM) overcome these classical method limitations and convenient tool to achieve an optimized condition for effective dye removal process. Furthermore, it helps to understand the statistical models, interaction with media parameters at varying levels in dye deolorization (Mohana et al., 2008; Pillai, 2017). RSM is an effective optimization tool; it concurrently predicts the factors and their interactions in few experimental trials. Karthikeyan et al. (2010) reported that optimization of culture conditions for dye decolorization by *Aspergillus niger* HM11 using response surface methodology. Similarly, response surface methodology was studied in the optimization of azo dye amido black 10B decolorization by *Kocuria kristinae* RC3 (Uppala et al., 2018)

Recent studies have reported that *Bacillus firmus* that was isolated from local sewage and textile waste water had up to 97-98% dye decolorization capacity (Arora et al., 2007; Ogugbue et al., 2011). Therefore, Remazol Golden yellow (RNL), an azo dye, which is widely used colorant in textile dyeing and printing industries of study area was used as model dye substance, to evaluate the decolorization efficiency of native bacterium. The present study was aimed to investigate the ability of *Bacillus firmus* (TSL9) strain to decolorize remazol golden yellow azo dye and also to optimize the conditions of various parameters in order to achieve maximum dye decolorization by using statistical tools such as Plackett-Burman and central composite design.

MATERIALS AND METHODS

Isolation and Screening of Dye Decolorizing Bacteria

Activated sludge was collected from a textile industry in Tiruppur, India and processed for isolation of native decolorizing bacterial strains. Pour plate method was employed on nutrient agar containing (g l^{-1}): 5g of peptone, 5g of sodium chloride, 3g of yeast extract, 3g of beef extract and 20g of agar. Consequently inoculated plates were incubated at

35 and 45 °C for 24-48 hours incubation intended for the isolation of mesophilic and thermotolerant bacteria respectively. Morphologically distinct colonies were isolated, and experimented for azo dye decolorization ability on Luria Bertani agar plates containing (g l^{-1}): 10g of casein enzymic hydrolysate, 5g of yeast extract, 10g of sodium chloride and 15g of agar amended with various concentrations (50, 100, 150, 200 and 250 mg l^{-1}) of remazol golden yellow dye (RNL). Spot inoculated plates were subsequently incubated at 37 °C for 4 days. The bacterial isolates showing clear zone in all the concentrations of dye were preferred as potential strain and used for further studies. The strain *Bacillus firmus* (TSL9) was found to be the most potential strain since it showed significant dye decolorizing efficiency in the agar plate.

Identification of Dye Decolorizing Bacteria

The selected decolorizing bacterium was identified by 16S rRNA sequencing method. Extraction of genomic DNA, PCR amplification and 16S rRNA sequencing were carried out in Xcelris Labs Ltd, Ahmedabad, India. The 16S rRNA consensus sequence analysis and alignment of closely related species were executed in BLASTN (www.ncbi.nlm.nih.gov/BLAST) and multiple alignment program (Clustal W) respectively. A Phylogenetic tree was constructed in MEGA version 5.0 via Neighbor Joining (NJ) method. The nucleotide sequences were deposited in NCBI GenBank with an accession number: JX316004.

Effect of Co-substrates on Azo Dye Decolorization

The presence and absence of nourishments on azo dye decolorization with bacterial strain TSL9 was studied in diverse broth media. About 100 ml of Luria Bertani broth containing (g l^{-1}): 10g of casein enzymic hydrolysate, 5g of yeast extract, 10g of sodium chloride**;** Yeast Extract broth containing (g l^{-1}): 5g of yeast extract, 5g of sodium chloride**;** and Bushnell Hass broth containing (g l^{-1}): 0.2g of magnesium sulphate, 1g of di-potassium hydrogen phosphate, 0.02g of calcium chloride, 0.05g of ferric chloride and 1g of ammonium nitrate was prepared and amended with 100 mg.l^{-1} concentration of remazol golden yellow dye. About 1% aliquot of 18 h bacterial culture (TSL9) was inoculated in these flasks and incubated at 37 °C for 3 days under a static condition. Samples were withdrawn aseptically at periodic intervals and centrifuged at 5000 rpm for 20 minutes. Absorbance (OD) of the cell free supernatant was recorded using UV spectrophotometer (Model: Cyberlab UV-100 USA) at 412 nm for remazol golden yellow dye. Decolorization efficiency was experimented in triplicates with abiotic control and the decolorization percentage (D %) was determined by the Equation (1), where A1 and A2 are the initial and final absorbance value.

$$D\,(\%) = \frac{A1 - A2}{A1} \times 100 \qquad [1]$$

Optimization of Dye Decoloriztion by Design of Experiments (DOEs)

Design of experiments (DOEs) is a statistical tool that exposes the individual and cumulative effects of factors that are involved in the dye decolorization by screening up to 'n-1' variables in just 'n' number of trails. Regression coefficients of response and fittest model prediction were analyzed with statistical software Minitab Version 15 (Tripathi & Srivastava, 2012).

Plackett-Burman Design for Screening of Significant Factors

The 2 k-factorial Plackett-Burman design of 12 trials were used to screen the significant factors of remazol golden yellow decolorization by *Bacillus firmus* (TSL9) and the effect of significant/ insignificant factors were studied. Screening of dye decolorization influencing factors and their actual values was shown in Table 1. Plackett-Burman design (Table 2) was generated on the basis of following first-order model equation.

$$Y= \beta_0 + \Sigma\, \beta_i X_i + \Sigma\, \beta_{ij} X_i^2 + \Sigma\, \beta_{ij} X_i X_i \qquad [2]$$

where, Y is decolorization, X_i is factor level, i is factor number, β_0 is model intercepts term, β_i *is* the linear effect, β_{ii} *is* squared effect, β_{ij} is interaction effect. Static decolorization experiments were performed in a basal broth medium, as per the statistical design with cell free control and the decolorization was assayed.

Table 1

Actual values of factors for Plackett-Burman design

Test variables	Lactose % (w/v)	Beef extract % (w/v)	pH	Temperature (°C)	Inoculums size % (v/v)	Dye concentration (mg/l)	Incubation period (hours)
Low level (-)	0.1	0.1	5	30	5	100	24
High level (+)	1.0	1.0	9	45	10	300	72

Table 2

Plackett-Burman design for screening of significant factors

Run order	LA % (w/v)	BE % (w/v)	pH	Temp (°C)	IS % (v/v)	DC (mg/l)	IP (hours)	DV-1	DV-2
1	1.0	0.1	9	30	5	100	72	1	1
2	1.0	1.0	5	45	5	100	24	1	1
3	0.1	1.0	9	30	10	100	24	-1	1
4	1.0	0.1	9	45	5	300	24	-1	-1

Table 2 *(Continued)*

Run order	LA % (w/v)	BE % (w/v)	pH	Temp (°C)	IS % (v/v)	DC (mg/l)	IP (hours)	DV-1	DV-2
5	1.0	1.0	5	45	10	100	72	-1	-1
6	1.0	1.0	9	30	10	300	24	1	-1
7	0.1	1.0	9	45	5	300	72	-1	1
8	0.1	0.1	9	45	10	100	72	1	-1
9	0.1	0.1	5	45	10	300	24	1	1
10	1.0	0.1	5	30	10	300	72	-1	1
11	0.1	1.0	5	30	5	300	72	1	-1
12	0.1	0.1	5	30	5	100	24	-1	-1

Where, LA - Lactose; BE - Beef Extract, Temp - Temperature, IS - Inoculum size, DC – Dye Concentration, IP - Incubation Period, DV1 & DV2 - Dummy variable, +1 denoted for high concentration; -1 denoted for low concentration

Significant Factors Optimization using Response Surface Methodology (RSM)

Optimization of the significant factors that resulted in the Plackett-Burman design was used for the maximum decolorization of remazol golden yellow by TSL9 by using response surface methodology (RSM). A central composite design (CCD) was applied to optimize the screened significant factors for effective dye removal. Significant factors, beef extract, inoculums size, and pH, were studied at five coded levels, -α, -1, 0, +1 and +α (Table 3). The actual values of these factors were calculated by following equation (Paul et al., 1992). Where $\alpha = 2^{n/3}$; here "n" was the number of factors and "0" was the central point.

$$\text{Coded value} = \frac{\text{Actual value - (high level + low level) / 2}}{\text{(High level - low level) / 2}} \quad [3]$$

The full factorial experimental runs were performed in triplicates as per the central composite design shown in Table 4, other than significant factors (sodium chloride 0.50% (w/v), dye concentration 200 mg/l, lactose 0.55% (w/v), temperature 37 °C and incubation period (48 hours) were kept constant. The correlation of significant factors with decolorization in central composite design can be described by following polynomial model shown in the equation.

$$Y = \beta_0 + \beta_1X_1 + \beta_2X_2 + \beta_3X_3 + \beta_{11}X_1^2 + \beta_{22}X_2^2 + \beta_{33}X_3^2 + \beta_{12}X_1X_2 + \beta_{13}X_1X_3 \quad [4]$$

Where Y is the response, β_0 is the intercept term, β_1, β_2 and β_3 are linear coefficient of significant factors, β_{11}, β_{22}, β_{33} are quadratic coefficient, β_{12}, β_{13}, β_{23} are interaction coefficient and X_1, X_2 , X_3, X_4 are the coded factors. The most fitting model for an effective response was found by determination of coefficient R^2 (Zhao et al., 2010).

With the support of response optimizer, the optimized values of significant factors were analyzed in statistical software Minitab Version 15. Optimal level of significant factors for maximum decolorization of remazol golden yellow by *Bacillus firmus* (TSL9) was verified with validation experiments by comparing the experimental and predicted values of decolorization percentage.

Table 3

Actual values of the significant factors for CCD

Variables	Unit	Five levels of variables				
		-α (-1.68179)	-1	0	1	+α (+1.68179)
Beef extract	% (w/v)	-0.20681	0.1	0.55	1	1.306807
Inoculums size	% (v/v)	3.295518	5	7.5	10	11.70448
pH	-	3.636414	5	7	9	10.36359

Table 4

Central composite design for optimization of decolorization with significant factors

Run Order	Pt Type	Blocks	Beef extract % (w/v)	Inoculums size % (v/v)	Ph
1	1	1	0.1	5	5
2	1	1	1	5	5
3	1	1	0.1	10	5
4	1	1	1	10	5
5	1	1	0.1	5	9
6	1	1	1	5	9
7	1	1	0.1	10	9
8	1	1	1	10	9
9	-1	1	-0.20	7.5	7
10	-1	1	1.31	7.5	7
11	-1	1	0.55	3.29	7
12	-1	1	0.55	11.71	7
13	-1	1	0.55	7.5	3.63
14	-1	1	0.55	7.5	10.36
15	0	1	0.55	7.5	7
16	0	1	0.55	7.5	7
17	0	1	0.55	7.5	7
18	0	1	0.55	7.5	7
19	0	1	0.55	7.5	7
20	0	1	0.55	7.5	7

RESULTS AND DISCUSSION

Screening of Dye Decolorizing Bacteria

About 20 bacterial strains were isolated from textile activated sludge and screened for ability of textile dye decolorization. These bacterial strains showed a wide variation of clear zones at different concentrations (50-250 mg/l) of remazol golden yellow dye amended in agar plates. Among the bacterial strains tested, TSL9 was found to be the most effective dye-decolorizing bacteria when subjected to further investigations. The efficiency of textile dye decolorization depends on the metabolic activity of microorganisms and their flexibility with the recalcitrant nature of dye (Khan et al., 2012). The bacterium TSL9 decolorized a mono azo dye, remazol golden yellow completely in all the concentrations that were incorporated.

Identification of Dye Decolorizing Bacteria

The dye decolorizing bacterial strain was identified by 16S rRNA sequencing method. Analysis of evolutionary history revealed that the sequence of TSL9 was closely associated with the diverse spectrum of the genus *Bacillus* depicted in the phylogenetic tree (Figure 1). The dye decolorizing bacterium was confirmed as *Bacillus firmus* (TSL9) and their sequence was submitted in the genbank database with the accession number of JX316004. In earlier studies, several *Bacillus* sp. have been reported for decolorization of textile dyes. Liao et al. (2013) reported that *Bacillus cereus* HJ-1 strain from azo dye contaminated river sediment is capable of decolorizing reactive black B (RBB).

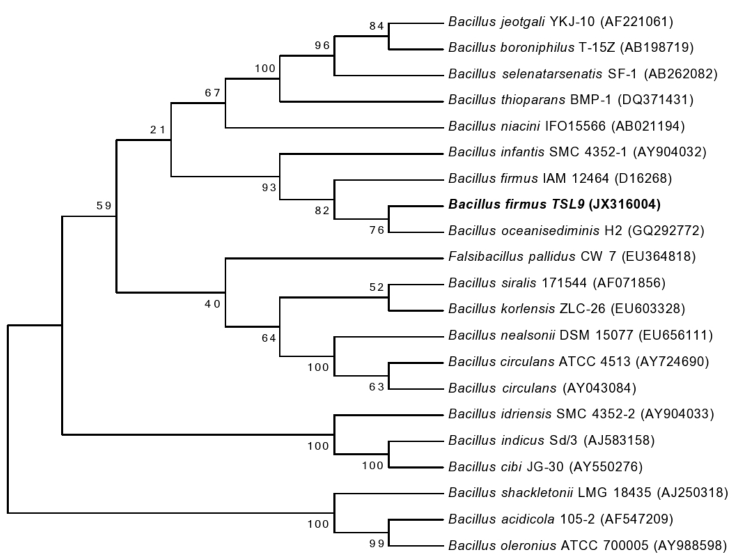


Figure 1. Phylogenetic tree of *Bacillus firmus* (TSL9) showing relationship between selected bacterial strains. Percentage numbers at the nodes indicate the levels of bootstrap support based on neighbour-joining analyses of 1000 replicates. Brackets represent the sequence accession numbers

Effect of Co-Substrates in Azo Dye Decolorization

Bacillus firmus (TSL9) was subjected to further experimentation by investigating the ability of dye removal in aqueous medium with/without carbon and nitrogen sources. The bacterial culture in Luria Bertani broth showed 69.78% remazol golden yellow removal after 24 hours, and 82.55% followed by 48 hours of incubation. After 72 hours, the rate of decolorization reached 97.23%. Carbon and nitrogen sources have different impacts on dye removal using micro-organisms; these elements may either be utilized for growth or may acted as an electron donor for dye reduction (Gonzalez-Gutierrez & Escamilla-Silva, 2009). In this study, the bacterial strain TSL9 accomplished maximal decolorization within a short period of incubation in Luria Bertani broth. This was mainly due to the amendment of casein enzymic hydrolyzate as a vitamin source for growth and metabolic activities. In contrast, yeast extract broth culture reached 27.9% dye removal with 24 hours, 62.03 and 73.81% decolorization were observed after 48 and 72 hours incubation respectively. Beyond 72 hours of incubation, dye removal efficiency did not increase. Addition of organic nitrogen sources, yeast extract and peptone in the medium enhanced the removal of dye as they act as NADH electron donor (Chang et al., 2000). In the case of Bushnell Haas broth, the strain TSL9 did not decolorize the dye significantly even at extended incubation period of 72 hours. When the carbon and nitrogen sources were absent in the basal broth medium, the bacterial cells failed to multiply. This can be consequently correlated to the absence of azo dye-catabolizing enzyme activity in the broth. These oxidoreductase enzymes such as azoreductase, laccase, and peroxidase are responsible for azo dye decolorization process (Mahmood et al., 2016). Due to deficiency of carbon and nitrogen compounds in dye structure, microbial cells require added nutritional factors to mediate the decolorization (Pillai, 2017; Sani & Banerjee 1999).

Optimization of Dye Decoloriztion by Plackett-Burman Design

Plackett-Burman design (PBD) results are shown in Table 5. There was a wide variation in remazol golden yellow decolorization using *Bacillus firmus* (TSL9), it ranged from 2.07 to 52.26%. The statistical analyses of estimated effects and regression coefficients indicate that a higher concentration of beef extract, incubation period, pH and temperature had a positive influence on decolorization. The reduction of the azo bond increased with higher levels of pH and led to the formation of basic aromatic amine metabolites (Willmott 1997). It was observed that maximum dye removal was obtained under optimal temperature, as it was favorable for azo reductase enzyme production (Chang et al., 2001). Consistently, extended incubation (70 hours) with pure culture of *Pseudomonas* sp. favored Reactive Blue 13 decolorization (Lin et al., 2010). On the contrary, lactose and size of the inoculums revealed the negative influence in lower concentration as shown Table 6. Bacterial cells did not prefer lactose as they assimilated dye as a carbon source (Bhavsar et al., 2018; Saratale et al., 2009) A vast volume of microbial cells is desirable for decolorization, owing to the toxic nature of dye molecule (Tan et al., 2013).

The most fitting model was confirmed by the determination of correlation coefficient (R^2) ($R^2 = 0.8975$) nearer to 1 for the regression equation (5), which means that the model can express up to 89.75% variation in the experiment. The value of the adjusted determination coefficient (adj R^2 = 71.81) corresponds to high significance of the model. Analysis of variance of the model is represented in Table 7. The F value of dye decolorization is 4.42. It was demonstrated that the model term was significant, the value of probability (P) > F is less than 0.085. The factors whose P-value were less than 0.10, indicated that the factors were highly significant (Khelifi et al., 2012).

Table 5

Plackett-Burman design of experimental and predicted decolorization (%) results

Run order	LA % (w/v)	BE % (w/v)	pH	Temp (^{o}C)	IS % (v/v)	DC (mg/l)	IP (hours)	Percentage decolorization	
								Experimental	Predicted
1	1.0	0.1	9	30	5	100	72	17.28	16.6
2	1.0	1.0	5	45	5	100	24	14.90	17.14
3	0.1	1.0	9	30	10	100	24	16.00	15.87
4	1.0	0.1	9	45	5	300	24	20.63	23.66
5	1.0	1.0	5	45	10	100	72	10.12	7.87
6	1.0	1.0	9	30	10	300	24	20.47	20.6
7	0.1	1.0	9	45	5	300	72	52.26	44.33
8	0.1	0.1	9	45	10	100	72	4.24	9.66
9	0.1	0.1	5	45	10	300	24	2.27	1.75
10	1.0	0.1	5	30	10	300	72	2.07	0.58
11	0.1	1.0	5	30	5	300	72	20.24	29.35
12	0.1	0.1	5	30	5	100	24	8.99	3.95

Where, LA - Lactose; BE - Beef Extract, Temp - Temperature, IS - Inoculum size, DC – Dye Concentration, IP - Incubation Period

Table 6

Statistical analysis of Plackett-Burman design

S. No	Variables	Effect	Coef	SE Coef	T	P
1	Constant		15.852	2.064	7.68	0.002*
2	Lactose	-3.270	-1.635	2.064	-0.79	0.473
3	Beef extract	13.353	6.677	2.064	3.23	0.032*
4	pH	11.873	5.937	2.064	2.88	0.045*

Table 6 *(Continued)*

S. No	Variables	Effect	Coef	SE Coef	T	P
5	Temperature	3.107	1.553	2.064	0.75	0.494
6	Inoculum size	-13.310	-6.655	2.064	-3.22	0.031*
7	Dye concentration	8.000	4.000	2.064	1.94	0.125
8	Incubation time	4.043	2.022	2.064	0.98	0.383
	R-Sq = 89.75% R-Sq (adj) =71.81%					

*Significant (p<0.05)

Table 7
ANOVA for Plackett Burman design

S. No	Source	DF	Seq SS	Adj SS	Adj MS	F	P
1	Main effects	7	1754.20	1754.20	1754.20	4.42	0.085
2	Lactose	1	28.61	28.61	28.61	0.50	0.517
3	Beef extract	1	513.65	513.65	513.65	9.06	0.040*
4	pH	1	435.49	435.49	435.49	7.68	0.050*
5	Temperature	1	31.27	31.27	31.27	0.55	0.499
6	Inoculum size	1	521.80	521.80	521.80	9.20	0.039*
7	Dye concentration	1	179.49	179.49	179.49	3.16	0.150
8	Incubation period	1	43.89	43.89	43.89	0.77	0.429
9	Residual error	4	226.87	226.87	56.72	-	-
	Total	11	1981.07				

*Significant (p<0.05)

$$Y = 15.852 - 1.635 \times \text{lactose} + 6.677 \times \text{beef extract} + 5.937 \times \text{pH} + 1.553 \times \text{temperature} - 6.655 \times \text{inoculum size} + 4.000 \times \text{dye concentration} + 2.022 \times \text{incubation period} \quad [5]$$

Standardized effects in a Pareto chart demonstrated the significant factors as shown in Figure 2. According to their statistical significance (p=0.10), beef extract, inoculum size and pH were found to be most important variables that influence dye decolorization by *Bacillus firmus* (TSL9).

Central Composite Design

The significant factors (beef extract, inoculums size and pH) and their mutual interactions were studied using response surface methodology (RSM), for maximum removal of remazol golden yellow dye using *Bacillus firmus* (TSL9). According to a central composite design (CCD), the decolorization output varied from 3.98 to 84.55% as shown in Table 8.

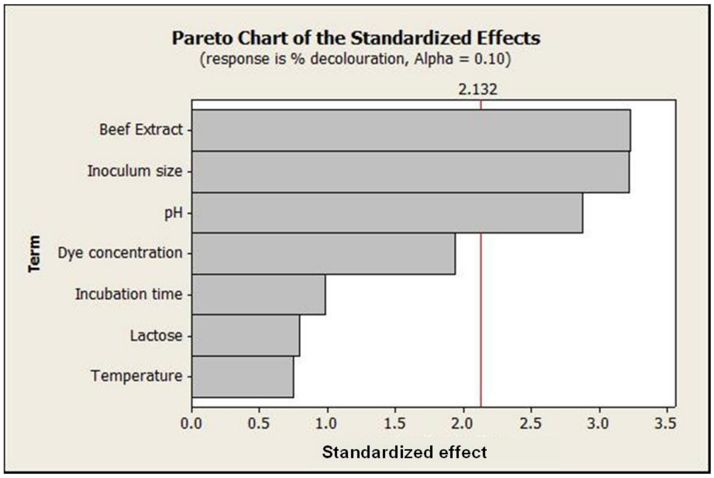


Figure 2. Effects of factors on dye decolorization in Pareto chart

Second order polynomial model equation (6) was fitted with the results of central composite design to reveal the dependence of the response. Smaller p-values with larger t-values related to coefficient value of factors were considered significant with a confidence level of 95% (Table 9). Correlation coefficient (R^2) was determined by regression analysis and it was found to be 98.68%, thus higher value of R^2 delivered presence of a good correlation between the factors and response (Zhao et al., 2010).

$$Y= 81.944 - 3.022 \times X_1 + 2.941 \times X_2 + 1.105 \times X_3 - 10.870 \times X_1^2 - 16.398 \times X_2^2 - 28.157 \times X_3^2 + 1.197 \times X_1 \times X_2 + 15.862 \times X_1 \times X_3 + 1.023 \times X_2 \times X_3 \quad [6]$$

Where Y is predicted response of dye decolorization (%), X_1, X_2 and X_3 were the coded values of significant factors.

In analysis of variance (Table 10) determined data, the F-value = 82.83 and probability value P =0.0 indicates the model is highly significant to predict the results. The linear (p=0.048), quadratic (p=0.000) and interaction (p=0.000) effect of the factors found in enhanced response/decolorization. Response counter plots with various levels of significant factors as shown in Figure 3, which signifies the relationship and interactive effect of two factors on dye decolorization, whereas a third factor value was maintained as middle level.

Table 8

Central composite design of experimental and predicted decolorization (%) results

Trails	Beef extract (X_1)	Inoculums size (X_2)	pH (X_3)	Percentage decolorization		Residual
				Experimental	Predicted	
1	0.1	5	5	36.64	43.57	-6.93
2	1	5	5	3.986	3.41	0.57
3	0.1	10	5	42.91	45.01	-2.10

Table 8 *(Continued)*

Trails	Beef extract (X_1)	Inoculums size (X_2)	pH (X_3)	Percentage decolorization		Residual
				Experimental	Predicted	
4	1	10	5	4.66	9.64	-4.98
5	0.1	5	9	12.71	12.01	0.69
6	1	5	9	33.12	35.30	-2.18
7	0.1	10	9	12.69	17.55	-4.86
8	1	10	9	48.27	45.62	2.64
9	-0.21	7.5	7	62.07	56.28	5.78
10	1.31	7.5	7	46.4	46.11	0.28
11	0.55	3.2	7	33.22	30.61	2.60
12	0.55	11.71	7	43.98	40.51	3.46
13	0.55	7.5	3.63	6.38	0.44	5.93
14	0.55	7.5	10.36	4.3	4.16	0.13
15	0.55	7.5	7	81.45	81.94	-0.49
16	0.55	7.5	7	84.55	81.94	2.60
17	0.55	7.5	7	78.16	81.94	-3.78
18	0.55	7.5	7	83.51	81.94	1.56
19	0.55	7.5	7	80.09	81.94	-1.85
20	0.55	7.5	7	82.86	81.94	0.91

Table 9

Estimated regression coefficients for Central composite design

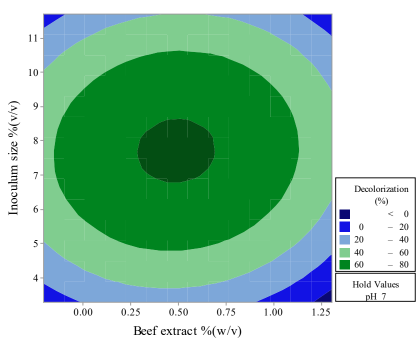
S. No	Variables	Coef	SE Coef	T	P
1	Constant	81.944	1.953	41.954	0.000*
2	Beef extract	-3.022	1.296	-2.332	0.042*
3	Inoculum size	2.941	1.296	2.270	0.047*
4	pH	1.105	1.296	0.853	0.414
5	Beef extract*Beef extract	-10.870	1.262	-8.616	0.000*
6	Inoculum size*Inoculum size	-16.398	1.262	-12.998	0.000*
7	pH*pH	-28.157	1.262	-22.320	0.000*
8	Beef extract*Inoculum size	1.197	1.693	0.707	0.496
9	Beef extract*pH	15.862	1.693	9.368	0.000*
10	Inoculum size*pH	1.023	1.693	0.604	0.559
	R-Sq = 98.68% R-Sq (adj) = 97.49%				

*Significant (p<0.05)

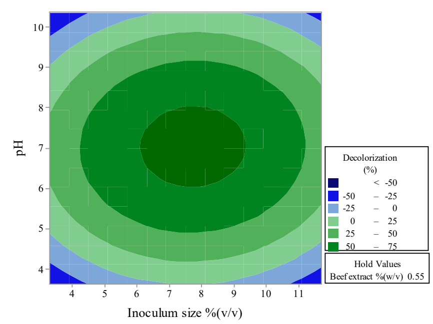
Table 10
ANOVA for central composite design

S. No	Source	DF	Seq SS	Adj SS	Adj MS	F	P
1	Regression	9	17097.8	17097.8	1899.75	82.83	0.000*
2	Linear	3	259.5	259.5	86.51	3.77	0.048*
3	Square	3	14805.6	14805.6	4935.21	215.19	0.000*
4	Interaction	3	2032.6	2032.6	677.53	29.54	0.000*
5	Residual error	10	229.3	229.3	22.93	-	-
6	Lack-of-fit	5	201.4	201.4	40.29	7.22	0.024*
7	Pure error	5	27.9	27.9	5.58	-	-
	Total	19	17327.1	-	-	-	-

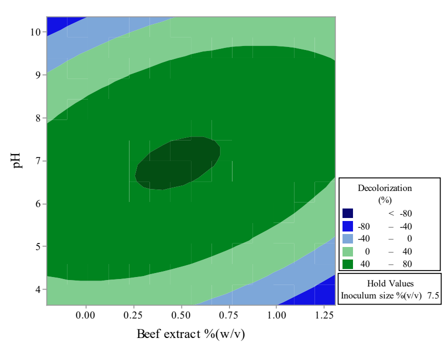
*Significant (p<0.05)



(a) Effect of beef extract and inoculum size



(b) Effect of pH and inoculum size



(c) Effect of pH and beef extract

Figure 3. Response contour plots between significant factors on decolorization (%)

The response optimizer tool was used to establish the most favorable point of significant factors for maximum dye removal rate, which was beef extract 0.48% (w/v), inoculum size 7.71% (v/v) and pH 6.96. Under the optimal level of significant factors, *Bacillus firmus* (TSL9) achieved 78.33% of decolorization, which agreed with the predicted decolorization (82.28%). The volume of inoculums provides sufficient biomass to perform decolorization of dye in an intermediate form. Similar to the present findings, beef extract 0.0025% (w/v) showed 97-98% decolorization of disperse dye by *Bacillus firmus* at a wide range of pH (6-9) (Arora et al., 2007). Ogugbue et al. (2011) reported the decolorizarion of Polar red B by *Bacillus firmus*, where maximum dye removal was obtained at pH 7-8. Bolstering the present investigation, Saraswathi et al. (2009) and Padhmavathy et al. (2003) accounted that *Bacillus firmus* acted as a potential candidate in the textile industry effluent treatment. Dafnopatidou and Lazaridis (2008) also stated that *Bacillus firmus* a halo tolerant strain had the ability to decolorize the dye at high salt concentration ranging from 0 to 60 mg/l.

CONCLUSION

The present investigation clearly demonstrated that the dye removal competence of *Bacillus firmus* (TSL9) can be related to the practical implementation constraints of textile industry. The dye decolorizing bacterial broth mediums require nutritional properties for their prospective decolorization and it was not able to decolorize remazol golden yellow dye in the absence of carbon and nitrogen sources. Application of statistical tool, Plackett-Burman design revealed the contribution of significant factors on dye removal using bacteria. Beef extract effectively utilized by the cells in the decolorization process than lactose as a carbon source in the medium. Thus, beef extract is suitable alternative organic nitrogen source which helps to avoid usage of inorganic nitrogen substances and generation of secondary pollutants usually occurred in the conventional effluent treatment at industrial level. Optimization of important factors for an effective RNL decolorization through Response surface methodology achieved a satisfactory response that supports in designing a cost effective process for dye decolorization. The optimized significant factors viz., beef extract, inoculum size and pH influenced the maximum dye removal efficiency in liquid broths, indicating the potential of native bacterium as an effective biodegradation tool for the decolorization of azo dyes from the effluents of the dye industries. Further scale-up studies are required on the decolorization of dye containing industrial effluents by isolated bacteria in order to assess the potential of *Bacillus firmus* as a prospective bioremediation agent.

ACKNOWLEDGEMENT

The authors are sincerely thankful to Vice Chancellor and Registrar of Periyar University, Salem for the financial support through University Research Fellowship (URF).

REFERENCES

Arora, S., Saini, H. S., & Singh, K. (2007). Decolorisation optimisation of a monoazo disperse dye with *Bacillus firmus*. Identification of a degradation product. *Coloration Technology, 123*(3), 184-190.

Arslan-Alaton, I., Gursoy, B. H., & Schmidt, J. (2008). Advanced oxidation of acid and reactive dyes: Effect of Fenton treatment on aerobic, anoxic and anaerobic processes. *Dyes and Pigments, 78*(2), 117-130.

Bhattacharya, S., Gupta, A. B., Gupta, A., & Pandey, A. (2018). Introduction to water remediation: importance and methods. In S. Bhattacharya, A. B. Gupta, A. Gupta & A. Pandey (Eds.), *Water remediation* (pp. 3-8). Singapore: Springer.

Bhavsar, S., Dudhagara, P., & Tank, S. (2018). R software package based statistical optimization of process components to simultaneously enhance the bacterial growth, laccase production and textile dye decolorization with cytotoxicity study. *PloS ONE, 13*(5), 1-18.

Chang, J. S., Kuo, T. S., & Chao, Y. P. (2000). Azo-dye decolorization with a mutant *Escherichia coli* strain. *Biotechnology Letters, 22*, 807-812.

Chang, J., Chou, C., Lin, Y., Lin, P., Ho, J., & Hu, T. L. (2001). Kinetic characteristics of bacterial azo-dye decolorization by *Pseudomonas luteola*. *Water Research, 35*(12), 2841-2850.

Dafnopatidou, E. K., & Lazaridis, N. K. (2008). Dyes removal from simulated and industrial textile effluents by dissolved-air and dispersed-air flotation techniques. *Industrial and Engineering Chemistry Research, 47*(15), 5594-5601.

Gonzalez-Gutierrez, L. V., & Escamilla-Silva, E. M. (2009). Reactive red azo dye degradation in a UASB bioreactor: Mechanism and kinetics. *Engineering in Life Sciences, 9*(4), 311-316.

Jadhav, S., Kalme, S. D., & Govindwar, S. P. (2008). Biodegradation of methyl red by *Galactomyces geotrichum* MTCC 1360. *International Biodeterioration and Biodegradation, 62*(2), 135-142.

Julkapli, N. M., Bagheri, S., & Hamid, S. B. A. (2014). Recent advances in heterogeneous photocatalytic decolorization of synthetic dyes. *The Scientific World Journal, 2014*, 1-25.

Kalyani, D. C., Telke, A. A., Govindwar, S. P., & Jadhav, J. P. (2009). Biodegradation and detoxification of reactive textile dye by isolated *Pseudomonas* sp. SUK1. *Water Environment Research, 81*(3), 298-307.

Karthikeyan, K., Nanthakumar, K., Shanthi, K., & Lakshmanaperumalsamy, P. (2010). Response surface methodology for optimization of culture conditions for dye decolorization by a fungus, *Aspergillus niger* HM11 isolated from dye affected soil. *Iranian Journal of Microbiology, 2*(4), 213-222.

Khan, R., Bhawana, P., & Fulekar, M. H. (2012). Microbial decolorization and degradation of synthetic dyes: a review. *Reviews in Environmental Science and Bio/Technology, 12*(1), 75-97.

Khehra, M., Saini, H., Sharma, D., Chadha, B., & Chimni, S. (2005). Decolorization of various azo dyes by bacterial consortium. *Dyes and Pigments, 67*(1), 55-61.

Khelifi, E., Touhami, Y., Thabet, O. B., Ayed, L., Bouallagui, H., Fardeau, M. L., & Hamdi, M. (2012). Exploring bioaugmentation strategies for the decolourization of textile wastewater using a two species consortium (*Bacillus cereus* and *Bacillus pumilus*) and characterization of produced metabolites. *Desalination and Water Treatment, 45*(1-3), 48-54.

Liao, C., Hung, C., & Chao, S. (2013). Decolorization of azo dye reactive black B by *Bacillus cereus* strain HJ-1. *Chemosphere, 90*(7), 2109-2114.

Lin, J., Zhang, X., Li, Z., & Lei, L. (2010). Biodegradation of reactive blue 13 in a two-stage anaerobic/aerobic fluidized beds system with a *Pseudomonas* sp. isolate. *Bioresource Technology, 101*(1), 34-40.

Liu, G., Zhou, J., Qu, Y., & Ma, X. (2006). Decolorization of sulfonated azo dyes with two photosynthetic bacterial strains and a genetically engineered *Escherichia coli* strain. *World Journal of Microbiology and Biotechnology, 23*(7), 931-937.

Mahmood, S., Khalid, A., Arshad, M., Mahmood, T., & Crowley, D. E. (2016). Detoxification of azo dyes by bacterial oxidoreductase enzymes. *Critical Reviews in Biotechnology, 36*(4), 639-651.

Mohana, S., Shrivastava, S., Divecha, J., & Madamwar, D. (2008). Response surface methodology for optimization of medium for decolorization of textile dye Direct Black 22 by a novel bacterial consortium. *Bioresource Technology, 99*(3), 562-569.

Mojsov, K. D., Andronikov, D., Janevski, A., Kuzelov, A., & Gaber, S. (2016). The applicationof enzymes for the removal of dyes from textile effluents. *Advanced Technologies, 5*, 81-86.

Mondal, P., Baksi, S., & Bose, D. (2017). Study of environmental issues in textile industries and recent wastewater treatment technology. *World Scientific News, 61*, 98-109.

Morales-Alvarez, E. D., Rivera-Hoyos, C. M., Poveda-Cuevas, S. A., Reyes-Guzman, E. A., Pedroza-Rodríguez, A. M., Reyes-Montano, E. A., & Poutou-Pinales, R. A. (2018). Malachite green and crystal violet decolorization by ganoderma lucidum and pleurotus ostreatus supernatant and by rGlLCC1 and rPOXA 1B concentrates: molecular docking analysis. *Applied Biochemistry and Biotechnology, 184*(3), 794-805.

Ogugbue, C., Sawidis, T., & Oranusi, N. (2011). Evaluation of colour removal in synthetic saline wastewater containing azo dyes using an immobilized halotolerant cell system. *Ecological Engineering, 37*(12), 2056-2060.

Padmavathy, S., Sandhya, S., Swaminathan, K., Subrahmanyam, Y. V., Chakrabarti, T., & Kaul, S. N. (2003). Aerobic decolorization of reactive azo-dyes in presence of various co-substrates. *Chemical Biochemical Engineering Q, 17*(2), 147-151.

Pillai, H. J. S. (2017). Optimization of process condition for effective degradation of azo bluedye by Streptomyces DJP 15. *Journal of Pure and Applied Microbiology, 11*(4), 1757-1765.

Rawat, D., Mishra, V., & Sharma, R. S. (2016). Detoxification of azo dyes in the context of environmental processes. *Chemosphere, 155*, 591-605.

Robinson, T., McMullan, G., Marchant, R., & Nigam, P. (2001). Remediation of dyes in textile effluent: a critical review on current treatment technologies with a proposed alternative. *Bioresource Technology, 77*(3), 247-255.

Sani, R. K., & Banerjee, U. C. (1999). Decolorization of triphenylmethane dyes and textile and dye-stuff effluent by *Kurthia* sp. *Enzyme and Microbial Technology, 24*(7), 433-437.

Saraswathi, K., & Balakumar, S. (2009). Biodecolourization of azo-dye (Pigmented red 208) using *Bacillus firmus* and *Bacillus laterosporus. Journal of Bioscience Technology, 1*(1), 1-7.

Saratale, R. G., Saratale, G. D., Chang, J. S., & Govindwar, S. P. (2011). Bacterial decolorization and degradation of azo-dyes: a review. *Journal of the Taiwan Institute of Chemical Engineers, 42*(1), 138-157.

Saratale, R. G., Saratale, G. D., Kalyani, D. C., Chang, J. S., & Govindwar, S. P. (2009). Enhanced decolorization and biodegradation of textile azo-dye Scarlet R by using developed microbial consortium-GR. *Bioresource Technology, 100*(9), 2493-2500.

Solis, M., Solis, A., Perez, H. I., Manjarrez, N., & Flores, M. (2012). Microbial decolouration of azo dyes: A review. *Process Biochemistry, 47*(12), 1723-1748.

Tan, L., Ning, S., Xia, H., & Sun, J. (2013). Aerobic decolorization and mineralization of azo dyes by a microbial community in the absence of an external carbon source. *International Biodeterioration and Biodegradation, 85*, 210-216.

Tony, B. D., Goyal, D., & Khanna, S. (2009). Decolorization of textile azo dyes by aerobic bacterial consortium. *International Biodeterioration and Biodegradation, 63*(4), 462-469.

Tripathi, A., & Srivastava, S. K. (2012). Biodegradation of orange G by a novel isolated bacterial strain *Bacillus megaterium* ITBHU01 using response surface methodology. *African Journal of Biotechnology, 11*(7), 1768-1781.

Uppala, R., Sundar, K., & Muthukumaran, A. (2018). Response surface methodology mediated optimization of decolorization of azo dye amido black 10B by *Kocuria kristinae* RC3. *International Journal of Environmental Science and Technology, 16*(8), 4203-4214.

Willmott, N. J. (1997). *The Use of Bacteria–Polymer Composites for the Removal of Colour from Reactive Dye Effluents* (PhD thesis). University of Leeds, United Kingdom.

Zhao, L., Zhou, J., Jia, Y., & Chen, J. (2010). Biodecolorization of acid red GR by a newly isolated *Dyella ginsengisoli* LA-4 using response surface methodology. *Journal of Hazardous Materials, 181*(1-3), 602-608.

Pertanika J. Sci. & Technol. 27 (4): 1881 - 1900 (2019)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Water Quality Effects on the Optimal Water Resources Operation in Great Karun River Basin

Iman Ebrahim Bakhsipoor, Seyed Mohammad Ashrafi* and Arash Adib

Department of Civil Engineering, Faculty of Engineering, Shahid Chamran University of Ahvaz, Ahvaz, Iran

ABSTRACT

The recent drastic decrease in the surface water resources quality has limited water resources managers in Great Karun river basin, southwest of Iran. In this research, the effects of water quality on the satisfaction of different demand sits in Great Karun river basin is modeled and studied by dint of systemic analysis principles based on the actual conditions in the river basin. In addition, different scenarios of water resources quality management are defined and the effect of implementing these scenarios on the demand satisfaction criteria is considered. The achieved results indicate the interactive relationship between quality conditions and performance of the system in demand satisfaction. It means that, the applied operating strategies can improve the demand coverage and reliability of the system only if the quality improvement is considered in extracting those strategies. The considered scenarios are able to improve average monthly demand coverage between 0% and 700% for Abadan city as a critical point within the system. Comparison of different quality management scenarios declares that decreasing the amount of agricultural demands has the most impact on reducing the flow contamination and improving the demand satisfaction throughout the basin, especially in downstream areas like Abadan City. This is due to the wide area of agricultural demands within the basin. Reduction of the agricultural demands causes least contaminated return flows on the one hand, and on the other hand it increases the amount of fresh water of river flows.

Keywords: Great Karun, model calibration, optimum water allocation, water quality modelling, water resources management

ARTICLE INFO

Article history:
Received: 11 January 2019
Accepted: 13 May 2019
Published: 21 October 2019

E-mail addresses:
iman.e.bakhshipoor@gmail.com (Iman Ebrahim Bakhsipoor)
ashrafi@scu.ac.ir; semo.ashrafi@gmail.com (Seyed Mohammad Ashrafi)
arashadib@scu.ac.ir (Arash Adib)
* Corresponding author

INTRODUCTION

As the rate of pollution in human societies increases due to industrialization and population growth, the violation of the flow

ISSN: 0128-7680
e-ISSN: 2231-8526

quality thresholds becomes a more important challenge to the water resource managers. As a result, the contamination/salinity level is the main limitation on water resources planning in many regions (Ashrafi & Mahmoudi, 2019; Karamouz et al., 2004; Mahmoud et al., 2011). In such conditions, it is not possible to manage a real-world water resources system through the water allocation modeling. Rather, the water quality issues must also be taken into account in the operating models. Therefore, the effect analysis of water quality on the allocation strategies of water resources based on the technical complexities of multi-reservoir problems (Ashrafi & Dariane, 2017) creates a fully complicated modeling problem, which calls for the involvement of special computational methods and powerful algorithms (Ghassemi et al., 1995).

In Iran, Great Karun River has the most discharge volume of water and is the longest river that flows into the Persian Gulf. The most important factors and challenges influencing the water resources quality of Karun River are the naturally saline branches and streams (Emamgholizadeh et al., 2014). In the Great Karun basin, the increased salinity of surface water has become an extremely big challenge for the water resources managers in this province since the construction of Gotvand Olya Dam and operation of inter-basin water transfer of Karun and Dez branches. Moreover, given the importance of the Great Karun River, numerous studies have been carried out on it, which have indicated the reduction of the quality criteria such as salinity and electrical conductivity indices in this river (Ehteshami et al., 2014; Hosseini-Zare et al., 2014; Jafarzadeh et al., 2004). The increased water demand in different sectors for purposes such as the expansion of agricultural activities, aquaculture farms, and urban development and industries can greatly influence the quality of water resources in Great Karun basin and the farming soil resources in the future. Due to the fertile soil and clement weather of this area, farming and industrial complexes have been created on the banks of Great Karun basin in addition to the traditional farms, which have a very long history in this region. Given the wide range of the agricultural, industrial, aquaculture, and urban development activities, the degree of contamination of the agricultural water resources and soil resources under study is increasing (Heydari et al., 2013). With a capacity of 4.5 billion cubic meters, Gotvand Olya is the second water reservoir in Iran by volume after Karkheh reservoir, and it is the biggest reservoir constructed over Karun River. As regards its position, this dam is located 10 km northeast of Gotvand City in Khuzestan Province and it is the most downstream reservoir dam over Karun River. The reservoir of this dam has been threatened with salinity since the beginning of its operation due to the placement of this dam in a salt formation (Radmanesh et al., 2013). In recent years (since 2 or 3 years after foundation of Gotvand Olya Dam), although the amount of water was adequate, the EC factor of river flows increased dramatically, several times. This has caused many expensive costs for users and stockholders of agriculture, aquaculture and municipal demands. Therefore, due to the special conditions of Khuzestan Province and the shortage of water resources in recent droughts, we need a management strategy to manage the associated basin, secure the highest level of allocation, and meet various demands with the minimum allowable quality. On the other hand, there is no an official Water Resources DSS for estimating the quality and quantity of available resources within the basin under different probable operating strategies.

Numerous studies have been carried out on the quality modeling and management of water resource systems (Gitau et al., 2016; Nikoo et al., 2013; Carmona et al., 2010; Estalaki et al., 2015; Kerachian & Karamouz, 2007). The main goal of most of these studies is achieving the best allocation strategy for quality thresholds satisfaction. In other words, water allocation must be determined such that it results in the highest level of quality defined in the system. Therefore, due to the special conditions of Khuzestan Province and the shortage of water resources in recent droughts, we need a management strategy to manage the associated basin, secure the highest level of allocation, and meet various demands with the minimum allowable quality (Moazami et al., 2016). Accordingly, in this research the demand satisfaction indices in Great Karun basin in Khuzestan Province is modeled considering the salinity of Karun and Dez rivers under different scenarios. This analysis was carried out by simulating the quality indices and quantity criteria in WEAP model to study the possibility of the maximum demand satisfaction in the presence of the maximum allowable salinity. Next, various scenarios were used to reduce salinity in different periods in Karun and Dez rivers within a 10-year period, and the effect of each scenario on the satisfaction of different demands in this basin was recorded.

STUDY AREA

The study area, with an area of 66930 km^2, included the Great Karun river basin (from the inlet to Khuzestan Plain (Gotvand City) to Abadan and Khorramshahr regions) as well as the Dez River from the inlet to Khuzestan Plain (Dez dam) to Band-e-Ghir (where it reaches the Great Karun River).

The Great Karun river basin includes the largest rivers by discharge volume of water (Karun, and Dez rivers) in Iran. This basin has important role in hydropower energy production and supplying drinking, industrial and agricultural demands in the provinces of Khuzestan, Chahar Mahal and Bakhtiari, Kohkiluyeh Boyerahmad and Lorestan. As regards the geographical coordinates of this basin, it is located at longitude 48^o 15’to 52^o 30’ E and latitudes 30^o 17’ to 33^o 49’N. Figure 1 represents the location of Great Karun river basin in Iran. Karun River originates from the Zagros Mountains 75 km south of Isfahan, and running to the north of Shushtar City, where it is divided into the Gargar and Shoteit branches. These two branches are merged in a region called “Band-e-Ghir” to form the Great Karun River along with Dez River. Smaller branches such as the Shur Dashte Bozorg, Balaroud, and Kohang branches also join these rivers, but they do not significantly change the trend of seasonal changes in the Great Karun River due to their annual input volumes. In addition, some of the pollutants flowing into Karun and Dez rivers include the agricultural effluent (from the irrigation networks, Sugar Cane development projects, and farming water rights), industrial sewage, and industrial effluents (Ashrafi & Mahmoudi, 2019).

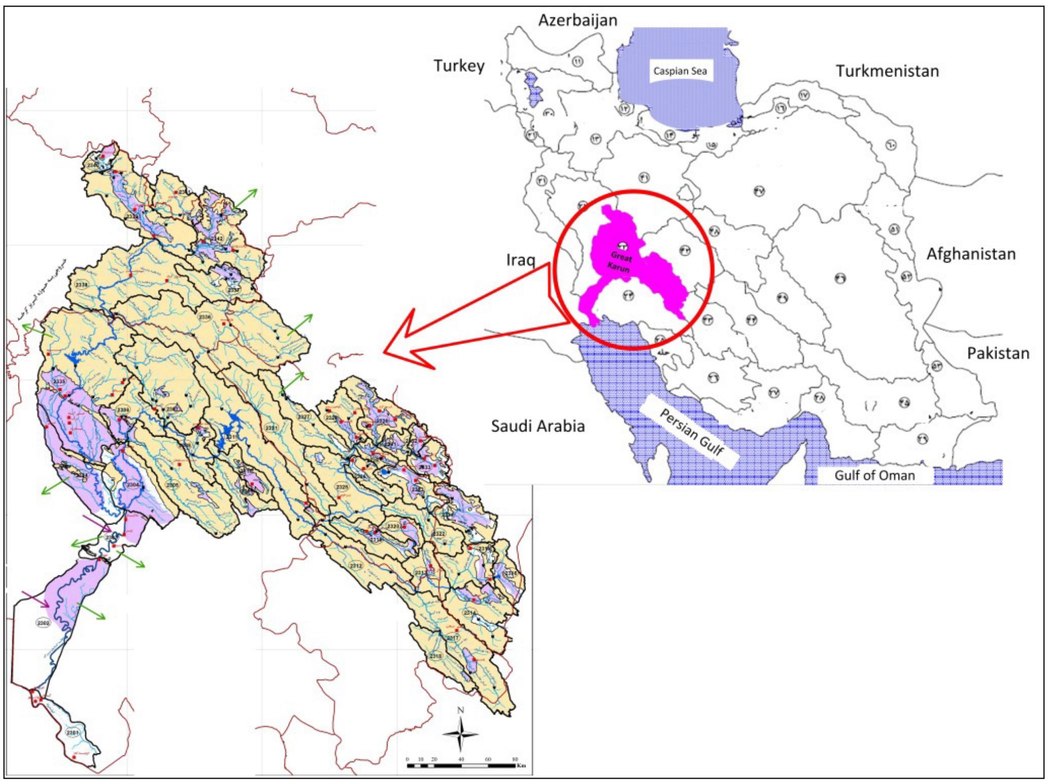


Figure 1. The location of Great Karun river basin in Iran

Based on the last study it should be noted that, the total water demands throughout the basin is growing up from 16700 *mcm/year* at near future, to 19000 *mcm/year* at long future. That indicates 14% growing rate for total demand (Dezab, 2011). About 74.6% of the basin consists of mountainous and highlands, while plains and low elevation areas cover about 25.4% of the basin. It is bounded by the Karkheh basin on the west, the Salt Lake, Gavkhouni and Bakhtegan-Maharlou on the north and east, the Zohreh-Jarahi Basin on the south and east. The basin joins the Arvand River and then the Persian Gulf at the outlet.

Seasonal and spatial variation of meteorological conditions in Great Karun river basin is significant. In flat areas, the summers are hot and humid, and the winters are moderate and slightly moist. At higher elevations, the winters are cold and dry and the summers are mild. Mean annual precipitation throughout the basin varies widely, ranging from 102 *mm* in South Ahvaz sub-basin to 1165 *mm* in Bazoft sub-basin. Generally, in mountain areas mean annual precipitation is slightly less than 715 *mm* which often occurs as snowfall during the mountain months. In flat Areas the precipitation occurs as rainfall. The actual evapotranspiration in mountain areas of the basin is around 12730 million cubic meters annually, which is almost equal to %35 of mountainous annual precipitation. In flat areas like South Ahvaz sub-basin, the annual evapotranspiration is up to %86 of annual precipitation.

MODEL DEVELOPMENT

In order to evaluate the effects of quality management strategies on the water allocation properties, the water resources simulation model is executed in different conditions. As presented in Figure 2, the configuration of Great Karun water resources system should be determined at the first step of the research. The system detail implemented within the configuration is really important, while simulation models are founded based on the system configuration.

The simulation models should be calibrated to evaluate the performance of the system accurately. The water resources management model is developed with monthly time step based on available data of the system. In this study, a simulation optimization approach is proposed to calibrate the simulation models. Water Evaluation and Planning (WAEP) model is applied for developing water resources simulation model while the Self-adaptive Melody Search (SaMeS) optimization algorithm is implemented as an optimization scheme.

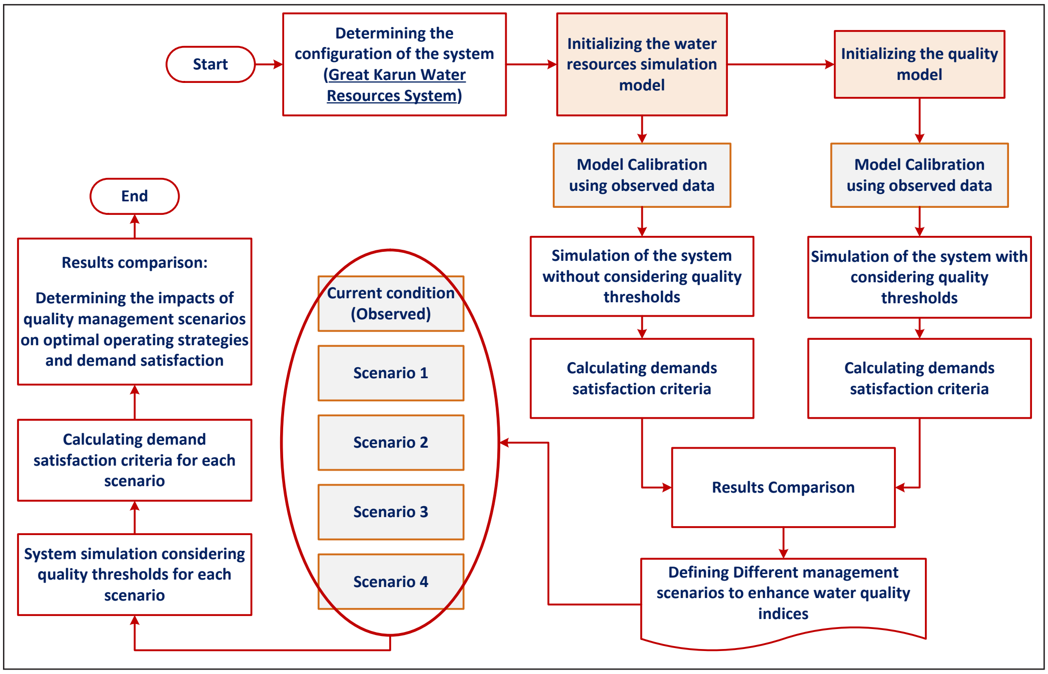


Figure 2. Flowchart of the presented research methodology

Results comparison of current condition evaluation can be really helpful while the effect of water quality on the system reliability is considerable. It is so important when the results of quality models usually are not considered by decision makers.

After establishing an accurate simulation model different strategies can be planned for enhancing the quality of Great Karun river flow. These strategies are implementable and based on the real condition of the system. The more probable strategies are modeled as four different quality management scenarios in this research. Evaluating the performance of the water resources system under these scenarios determines the effect of the water quality strategies on the allocation and system sustainability. Moreover, the most efficient strategy for reducing the flow contamination can be determined at the end.

Figure 3 presents the schematic of the Great Karun water resources system. Different distributed demands along the system are aggregated within some lumped demand nodes within the modeled basin in this study. Where, the drinking and municipal demands, the agricultural demands, Sugar Cane development sites, traditional farming water rights, aquaculture demands and industrial needs are considered as the main demand components of water resources system. Moreover, return flow of irrigation networks is modeled as a very effective factor in increasing the Electrical Conductivity (EC) indicator of river flows.

In this paper, the Water Evaluation and Planning system (WEAP) model is applied to simulate the water resources system of Great Karun river basin considering water quality and quantity issues under different scenarios. WEAP as a well-known generic tool for integrated simulation and analyzing water resources systems was developed by the Stockholm Environment Institute (SEI). WEAP integrates a spectrum of hydraulic-physical processes with demand management and installed infrastructure into a simple and comprehensive solution. In addition, it can be connected to MODFLOW, for accounting the surface water and groundwater interactions and to QUAL2K for tracing the water quality throughout an integrated system (da Silva & Alves, 2016).

The distributed agricultural and industrial demands are defined as integrated demand sites in three separated sub-areas of the basin, viz. Dez reservoir to Band-e-Ghir, Gotvand Olya reservoir to Band-e-Ghir, and Band-e-Ghir downstream domains (the Great Karun). In each sub-area, the corresponding municipal demands of big cities (e.g. Ahvaz, Abadan Khoramshahr, Dezful, Shushtar, Gotvand) are simulated via separate demand sites. Other drinking and municipal demands of small cities and villages are modeled as aggregated demand nodes in each sub-area. For more details about the developed simulation model refer to Ashrafi and Mahmoudi (2019).

To simulate the EC indicator within the considered system, the WEAP quality model is applied where the decay factor of EC parameter is assumed as zero. According to the National Drinking Water Standard of Iran, the EC values lower than 1650 μmho, values between 1650 and 2500 μmho, and values higher than 2500 μmho represent satisfactory

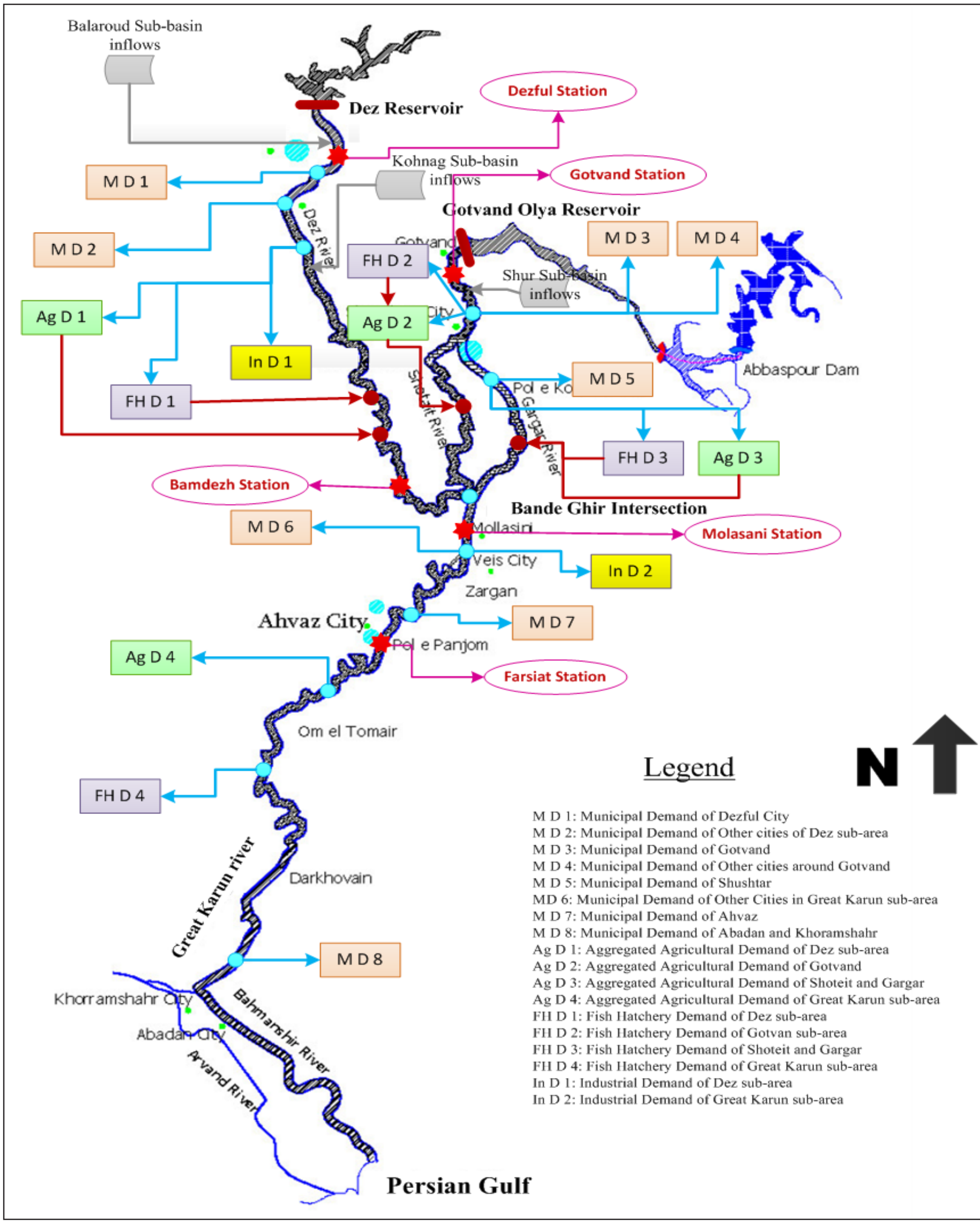


Figure 3. Schematic of the simulated Great Karun river basin

drinking water quality, allowable water quality, and non-allowable water quality, respectively. In this research, the EC of 2500 μmho is used as the minimum allowable water quality for municipal drinking water uses. Besides, the water allocation problem can be formulated through equation (1) considering water quality constraints. In other words, each demand is only met if the value of quality index does not exceed the allowable limit. As a result, it is possible to study the effect of violation of the salinity threshold on the satisfaction of different demand sites throughout the basin.

$$Max\ (\mathbb{M} = Demand\ Satisfaction\ Rate) \equiv Min\left(\mathbb{Z} = \left\|\overrightarrow{TD} - \vec{R}\right\|\right)$$

$$Subject\ to: \begin{cases} Physical\ and\ Conceptual\ Constraints\ of\ the\ System \\ Mass\ Balances\ for\ different\ Elemnts \\ Quality\ Constraints\ (EC_{sim}(i) \geq\ EC_{Per}(i)) \end{cases} \quad [1]$$

Where, M is a simulated value of objective function that is equal to the absolute differences of Target demands and water releases to supply demands. $\overrightarrow{TD}$ indicates the target demands as an $m \times t$ dimensional matrix, and $\vec{R}$ indicates the water releases which can be determined as a $m \times t$ dimensional matrix. The simulated value of EC parameter in the ith control node is formulated by $EC_{sim}(i)$ and $EC_{Per}(i)$ determines the permissible Electronic Conductivity at the same location. Where m, and t represents the number of demand sites, and the number of simulated time steps, respectively. And the satisfaction rate for i^{th} demand site at n^{th} time step is calculated as follows;

$$DC_{i,t} = \begin{cases} \dfrac{R_{i,t}}{TD_{i,t}}, & if\ R_{i,t} < TD_{i,t} \\ 1, & otherwise \end{cases} \quad [2]$$

Where, $DC_{i,t}$ stands for demand coverage of a certain demand site at a specific time step, which is identical to satisfaction rate.

CALIBRATION OF THE SIMULATION MODEL

Since modeling the flow quality parameters through the river basin is influenced by the distribution of the hydraulic and volumetric parameters of flow, the water resources system simulator must be calibrated before modeling EC distribution. In the water resources simulation models, model calibration refers to the preparation of the model for simulation of the real-world systems. This is carried out through the approximation of some model parameters based on observations and it can be explained as an inverse problem as follows (Sun & Sun, 2015);

$$\theta_{qs} = arg\underbrace{min}_{\theta}\left\|u_D(\theta) - u_D{}^{obs}\right\|, \qquad \theta \in P_{ad} \quad [3]$$

Where, θ_{qs} is a quasi-solution achieved by the optimization process, $u_D(\theta)$ indicates the simulation model output for the certain parameter set, θ, and $u_D{}^{obs}$ represents the observed value correspond to the current model output. It is worth noting that the water resources models cannot be expected to exactly achieve the same values as the observed data. This is why the imposed management strategy of the system is not necessarily based on the operational principles applied in the modeling. Moreover, many of the basic constraints of the real-world system may be neglected or simplified during the expansion of the computational model. A simulation-optimization approach can be implemented to solve equation (3) as an indirect method. Accordingly, Equation (4) states a composite mapping ($\mathcal{DM}$), which summarizes the forward model ($\mathcal{M}$) and the sampling mapping ($\mathcal{D}$) to obtain the system outputs corresponding to the observations using a specified parameter set (θ).

$$\mathcal{DM}(\theta) = u_D(\theta) \qquad [4]$$

Considering the simulation model structure, in this calibration we are searching to obtain proper values of water losses percentages in different sub-areas (δ_{Loss}), the agricultural and industrial return flows (ReF_{Ds}), and the demand supply priorities (Pr_{Ds}) for meeting all simulated demands. Hence, relation (4) can be rewritten as follows;

$$\mathcal{DM}(\theta) = \mathcal{DM}(\delta_{Loss}, ReF_{Ds}, Pr_{Ds}) \xRightarrow[\text{Simulation Model}]{\text{Through the}} u_D(\delta_{Loss}, ReF_{Ds}, Pr_{Ds}) \cong u_D{}^{obs} \qquad [5]$$

Finally, equation (1) can be regarded as an applied optimization model as follows;

$$Min\ RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}[(Q_{model})_i - (Q_{obs})_i]^2}$$

$$Subject\ to: \begin{cases} All\ Contstraints\ adopted\ within\ the\ Simulation\ model \\ All\ Constraints\ imposed\ by\ the\ operating\ Strategy \\ All\ Constraints\ derived\ from\ the\ system\ configuration \\ Q_{model} = u_D(\delta_{Loss}, ReF_{Ds}, Pr_{Ds}) \\ \Rightarrow\ u_D(\delta_{Loss}, ReF_{Ds}, Pr_{Ds}) = Q_{obs} \end{cases} \qquad [6]$$

The advantage of indirect methods in solving such inverse problems is that all of the optimization constraints (derived from the real-world conditions) can be accounted in the simulation model. Figure 4 presents the schematic of the calibration process of the developed water resources simulation model.

The Self-adaptive Melody Search (SaMeS) optimization algorithm was used in this research to find the optimal set of the model parameters. Melody Search algorithm is an algorithm that was introduced for the first time by Ashrafi and Darian (2011). The ability of this algorithm to solve different engineering problems has been proved (Ashrafi &

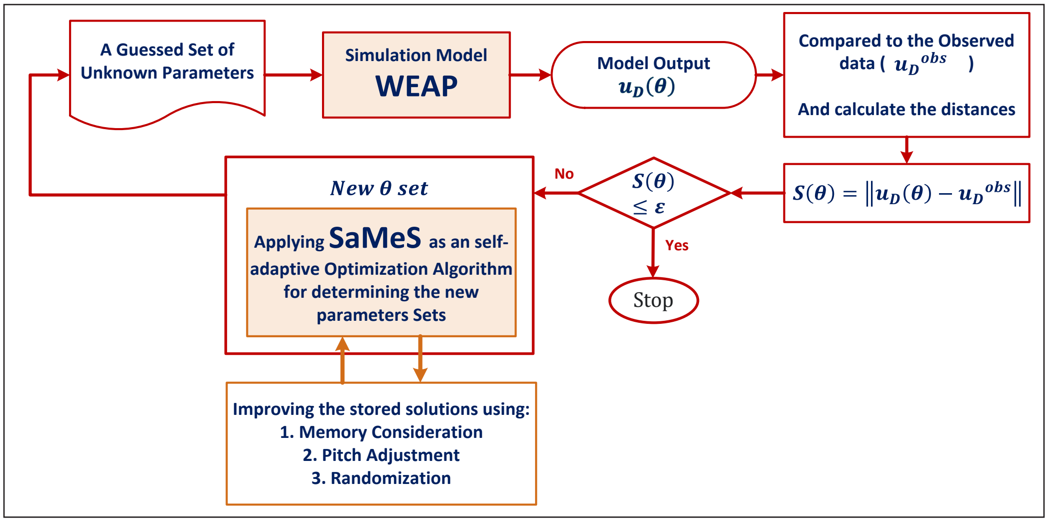


Figure 4. Schematic of the performed calibration process

Dariane, 2013). In addition, this algorithm is developed as a self-adaptive algorithm and it has been used numerously to optimize the water resources management problems as a powerful algorithm (Ashrafi & Kourabbaslou, 2015; Ashrafi et al., 2017; Ashrafi & Dariane, 2017). The three main operators used in this algorithm to find the optimal solution are the Memory Consideration, Pitch Adjusting, and Randomization operators. The modeling of different sub-memories interactions considerably enhances the algorithm performance. See Ashrafi and Kourabbaslou (2015) to find more details about MeS optimization algorithm. The historical observed data of 5 hydrometric stations in the Great Karun river basin are used to carry out the model calibration. These stations are Gotvand, Dezful, Bamdezh, Molasani, and Farsiat hydrometric stations. The locations of these hydrometric stations are shown in Figure 3. The result of calibration process is summarized in Table 1 while the Root Mean Square Error (*RMSE)* and R-Squared (R^2) indicators are presented for estimating the proximity of simulated and observed data. In all hydrometric stations discharge values of different time steps are simulated and compared to the observed ones. As shown in Table 1, the calibrated model can simulate the Great Karun water resources system accurately. Figure 5 shows the results comparison for all considered stations. The general R-Squared indicator is above 0.95 which indicates a good accuracy of the modeling.

Table 1
The results of calibrated model in distributed hydrometric stations

	Gotvand	Dezful	Bamdezh	Molasani	Farsiat
R^2	0.9372	0.9561	0.9178	0.97	0.9316
RMSE (mcm)	52.43	44.052	82.40	35.39	50.26

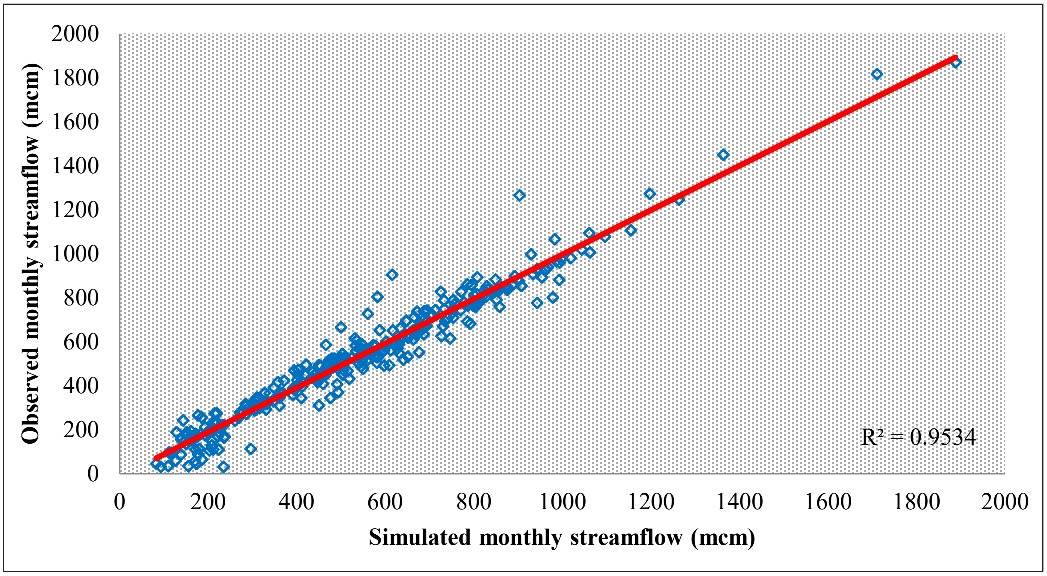


Figure 5. Overall comparison between simulated and observed stream flows in Gotvand, Dezful, Bamdezh, Molasani and Farsiat stations

CONSIDERED SCENARIOS

Water resources system modeling was carried out considering two basic scenarios and four probable scenarios. The basic scenarios are named Scenario 0 and Scenario 00. The probable scenarios were adopted as the proposed solutions to reduce water salinity and enhance the quality of supplied water for critical consumptions. The critical consumptions refer to the water used for drinking purposes in Ahwaz and Abadan Khoramshahr cities that are downstream the intersection of Dez and Karun rivers and are facing the over-salinity problem in some months.

Scenario 0

In this scenario water resources management modeling is carried out regardless the water quality. The results of this scenario form a basic insight about available resources within the system which is helpful to make a fair comparison and understand the effect of water quality on the demand satisfaction process.

Scenario 00

In this scenario the water allocation modeling is performed under water quality constraints for drinking and municipal water demands. The results of this scenario make a good understanding about the real condition of the Great Karun river basin.

Scenario 1

Since the observation data on the study area indicates that the agricultural effluent flowing into Dez and Karun rivers significantly influences the salinity of these rivers, it is possible to improve river water quality by reducing the water consumption in the agriculture sector by new irrigation methods. Hence, assuming a 50% decrease in water consumption using the new irrigation methods, the Dez and Karun sub-basins were modeled in WEAP. Equations 7 and 8 represent the variable changing in scenario 1.

$$WCon(i,t) = 0.5 \times BWCon(i,t) \quad [7]$$

$$ReF(i,t) = \left(1 - WCon(i,t)\right) \times WS(i,t) \quad [8]$$

where, $WCon$ indicates the amount of water consumption, $BWCon$ is the basic consumption rate, *ReF* stands for the amount of return flow, and *WS* is the amount of released water to supply demands. Equation (6) has been formulated for the i^{th} demand site at t^{th} time step. The basic water consumption for different demands is achieved from the observed data or calibration results. Other scenarios executed using the basic consumption values.

Scenario 2

The other solution for reducing the EC factor caused by the agricultural effluent is to treat it before it enters the river. To this end, a scenario was formulated for the construction of treatment plants to treat agricultural effluent in the sub-basins of Karun and Dez rivers and reduce the salinity of the returned flows by 50%.

$$EReF(i,t) = (1 + \alpha) \times EWS(i,t) \quad [9]$$

$$\beta_{S2} = 0.5 \times \alpha \quad [10]$$

$$EReF(i,t) = (1 + \beta_{S2}) \times EWS(i,t) \quad [11]$$

The agricultural impact on the EC factor of river flows is assumed linear (Equations 9 and 11).where, *EReF* is the EC factor of return flows, and *EWS* is the EC factor of released water to supply demands. The α and β_{S2} stand for linear coefficients for modelling the effects of agricultural impacts on EC factor without and with treatment, respectively.

Scenario 3

Since the EC of the water discharged from Dez reservoir is lower than that of the water released from Gotvand Olya reservoir, it determines the quality of the water provided to the Band-e-Ghir downstream region. However, with the transfer of high-quality water from upper Dez branches to other basins, the shares of Khuzestan and Dez rivers of this quality water have decreased dramatically, resulting in a considerable decrease in the quality of water in Dez and Karun rivers. Hence, in another scenario, the effect of reducing the inter-basin water transfer from Dez river branch is studied. To this end, it is assumed that the amount of transbasin water diversion from Dez branches is decreased to 400 mcm/year. It means that the annual transferred water from Dez River is relatively dropped by half. And hence, the effect of increasing the high quality water of Dez River on the demand satisfaction in downstream areas of the basin is considered.

Scenario 4

In this scenario, the effect of the simultaneous implementation of the three previous scenarios on the water quality improvement is studied.

RESULTS AND DESCUSSION

In the Scenario 0, water allocation was carried out regardless of the quality constraint. This scenario was defined to study the effect of system contamination on the satisfaction of demands. The focus of this research was mainly the degree of satisfaction of the drinking and municipal water demands. Figure 6 shows the average monthly satisfaction of different demands based on the last 10-years statistics.

In the simulation carried out for different municipal, agricultural, and industrial uses in Dez and Karun sub-areas, it was found out that the inputs of Dez and Karun rivers are enough for the quantitative satisfaction of all demands in all months and during most

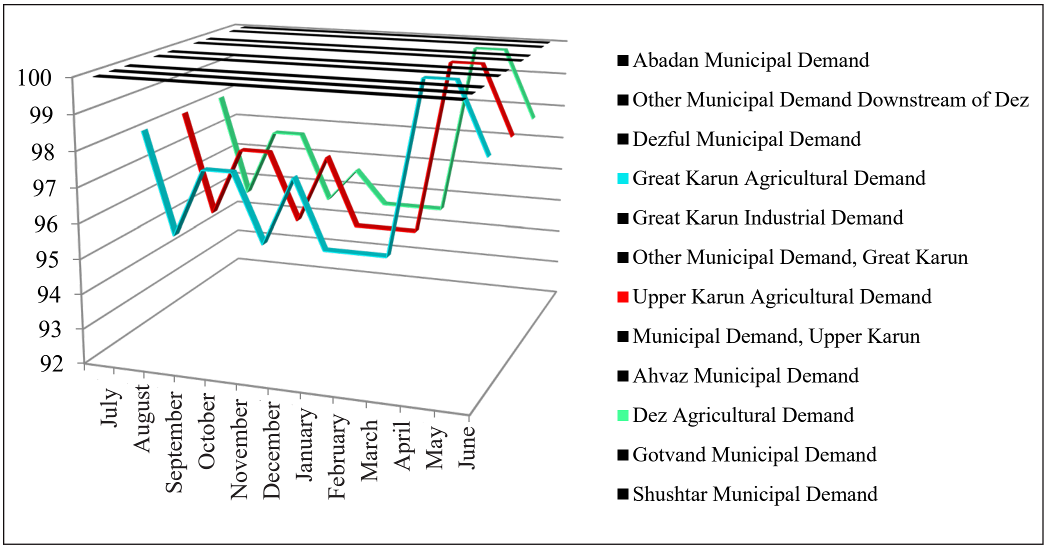


Figure 6. Averaged monthly demand coverage in Scenario 0

simulation years. There were only small shortages in meeting the agricultural demands, which had lower priorities than the municipal needs.

To evaluate the effect of water quality on the demand satisfaction, the simulation was carried out by adding a 2500 μmho quality constraint for the municipal demands. This simulation is named scenario 00. Based on the Iranian water quality standards (Torabian, & Shahavi, 2017) for municipal demands, and in order to maximize the use of available water resources, the water quality threshold is set to 2500 μmho in this study. This threshold is assigned for the quality of river flows with assumption of existing suitable pretreatment before municipal consumptions. The WEAP model does not allocate water to a demand unless the water quality at the point of survey does not violate the minimum limit, which is 2500 μmho. Figure 7 presents the simulation results under this scenario.

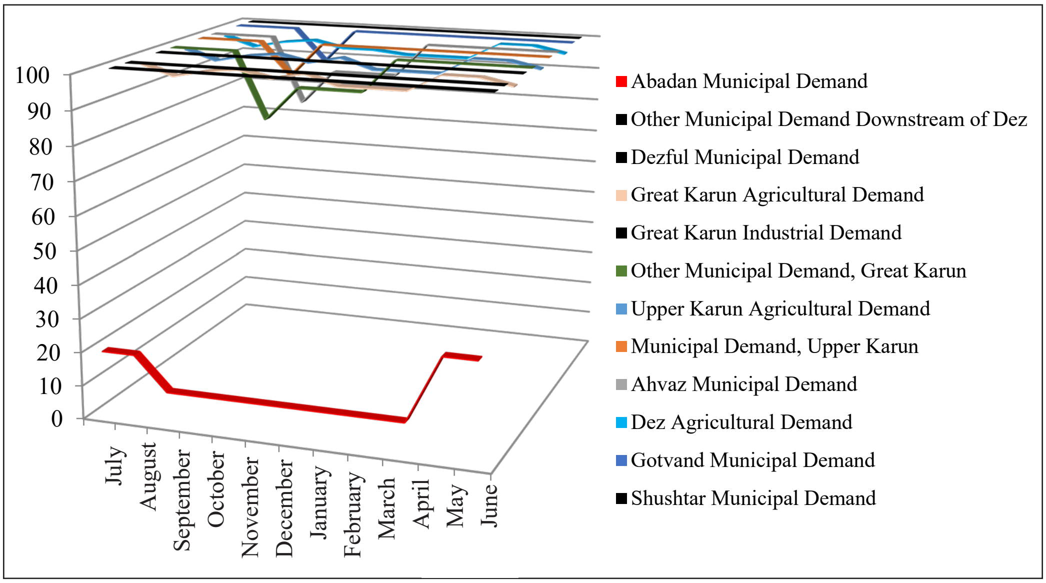


Figure 7. Averaged monthly demand coverage in Scenario 00

By adding adequate quality constraints for different demands, the minimum water quality limit is not met for the municipal water demand of Ahwaz in October, November, December, and January, as well as the municipal water demand of Abadan and Khoramshahr City in all months. Moreover, the municipal water demand of the other cities in the Great Karun sub-basin is not met during October, November, and December due to the high EC values. As seen, the drinking water demands of Gotvand City and other cities downstream Gotvand Olya reservoir, which are placed before the intersection of Karun and Dez rivers, is not met completely during October due to the high EC values of the water discharged from Gotvand Olya in some years of the simulation period.

Under scenario 1, a 50% decrease is observed in the water demands of the agriculture sector and the acceptable quality of the drinking water of Ahwaz City is secured in all months except for October and November. In Abadan, the supply of quality water in most months has increased considerably. Imposing this scenario enhances the monthly average coverage for Abadan municipal demand between 0% and 300% at different months, where the value of EC parameter is decreased between 13% and 29%. Therefore, it seems that reducing the agricultural needs can be a useful solution for improving water quality and providing water quality by reducing the volume of return flow to the river which often is highly unsuitable. However, it does not solve the problem completely.

Modeling under the scenario 2 secured the allowable quality of the water meeting the municipal water demands of Ahwaz City in all months except for October. However, the problem of meeting the municipal demand of Abadan City has not been solved yet and there is a considerable deficiency in this regard. Under this scenario, no improvement of monthly average demand coverage in Abadan city is observed. However, the average value of EC factor improved around 7%. It means that Scenario 2 is not useful in solving the problem of suppling Abadan municipal demand.

Scenario 3 reveals that reducing the amount of Dez River inter-basin water transfer to 400 *mcm/year* supplies the required quality of municipal demand in Ahwaz City in all months, while the municipal demands of Abadan City remains unmet in whole of the year. Under this scenario the minimum improvement of EC parameter in Abadan station is 11% in August and the maximum value of EC improved is 25%. The reduction of EC value causes the enhancement of monthly average coverage for Abadan municipal demand between 0% and 200% in different months.

In scenario 4, the municipal water demand of Ahwaz City is completely satisfied. It was also found out that the quality of water supplied to Abadan City is improved considerably in this scenario. The minimum and maximum improvement of EC factor is around 29% in November and 38% in January and April, respectively. However, the water demands of this city are not fully met in any month. Figure 8 shows the average percentage of the monthly supply of water to Abadan City under different scenarios. The minimum and maximum demand coverage enhancement in Abadan city under Scenario 4 is 150% and 700%, respectively. The minimum enhancement is achieved in May, and June and the maximum improvement is occurred in April, when the agricultural demands reach their maximum values.

As seen in Figure 8, scenario 4 has the highest impact on the increased supply of drinking water to Abadan City due to the concurrent implementation of the three quality improvement scenarios. As compared to the current situation (scenario 00), the implementation of scenario 2 cannot meet the drinking water quality threshold in Abadan City and fails to improve the satisfaction of its needs despite the increase in the satisfaction

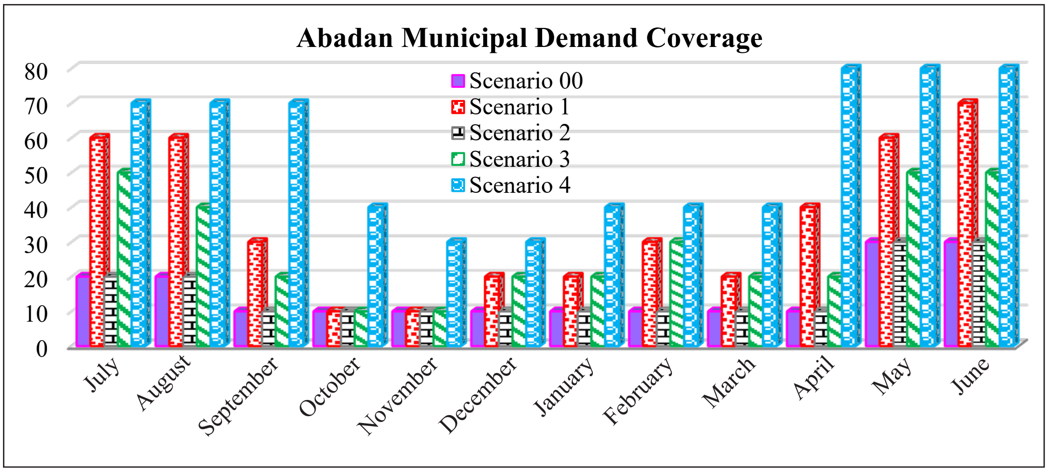


Figure 8. The average monthly demand coverage of Abadan city in different scenarios

of municipal water demands of Ahwaz City and the improvements in water quality in the point of supply in Abadan City. In addition, scenarios 1 and 3 have relatively similar effects on the satisfaction of the drinking water needs of Abadan City in all months of the year except for October, November and February when scenario 1 improved the demand satisfaction rate more. Figure 9 also shows the average monthly EC values of the flow in the simulation period, which confirms the previously mentioned finding.

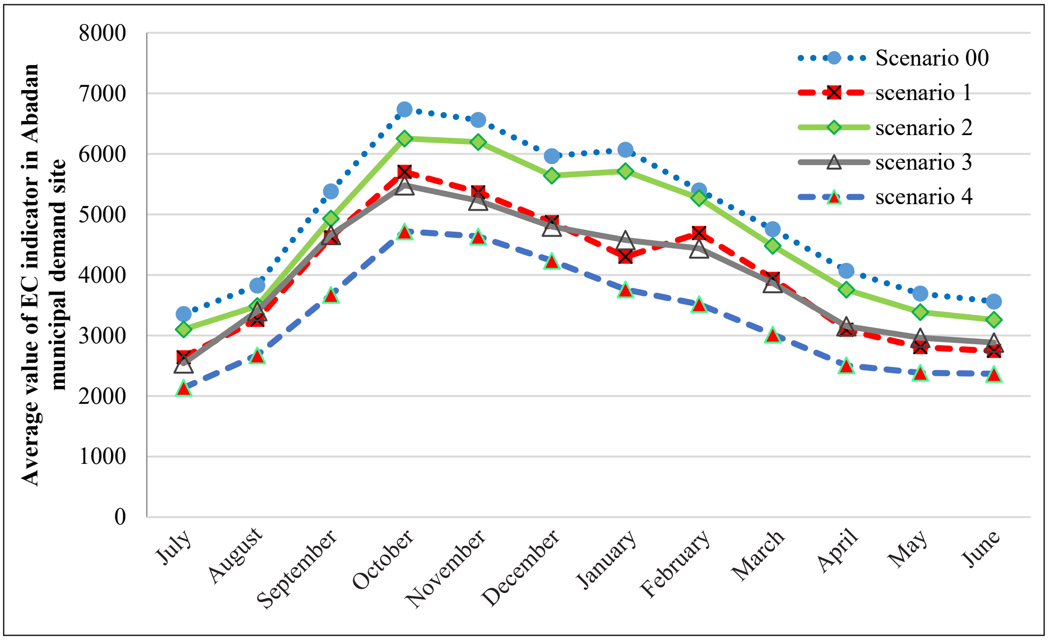


Figure 9. Average monthly value of EC of Great Karun river flow in Abadan municipal demand site

According to the basic concepts of water resources management, enhancing the demand coverage throughout the system results in improving system reliability in a certain time horizon (Hashimoto et al., 1982). It means that, the applied operating strategies which can improve the demand coverage, are able to enhance the reliability of the system.

CONCLUSION

In this research, the water quality of Dez and Karun rivers in Khuzestan was evaluated and the EC index was simulated in WEAP model to evaluate the effect of water quality of these rivers on the satisfaction of the municipal, agricultural, and industrial demands under different scenarios. The results of simulations in different cases indicate that the degree of demand satisfaction in this basin can be changed significantly by the quantitative parameters. In other words, despite the acceptable quantitative flow values the system managers may fail to meet the needs of a region. This has happened several times in recent years in the downstream of Great Karun basin. This is highly important in the disputes over the inter-basin transfer of water. That is to say, the shares of different basins should not be solely determined through quantity flow management as the water quality modeling must also be taken into account. According to the research results, under the existing conditions, it is not possible to fully meet the municipal water demands of the downstream Great Karun basin if the existing hydrologic conditions continue in the future. In this regard, reducing the agricultural water demands in the basin and reducing the withdrawal from Dez branches make the biggest contribution to the satisfaction of the drinking water demands in this basin. Therefore, a 50% decrease in the agricultural water demand, the construction of a treatment plant for treating the agricultural effluent in Dez river sub-area, and the decreased transfer of fresh water from Dez River to other watersheds will significantly improve the current conditions.

ACKNOWLEDGEMENT

The authors gratefully acknowledge all employees of the Environmental Office of the Khuzestan Water and Power Authority (KWPA) who provided insight and expertise that greatly assisted the research although they might not agree with all of the interpretations/ conclusions of this paper. Moreover, the authors are grateful to Dr. H. Hasounizadeh, Mr. E. Mostaghimzadeh, and Mrs. E. Saki for their assistance in gathering the research data.

REFERENCES

Ashrafi, S. M., & Mahmoudi, M. (2019). Developing a semi-distributed decision support system for great karun water resources system. *Journal of Applied Research in Water and Wastewater, 6*(1), 16-24.

Ashrafi, S. M., & Dariane, A. B. (2011, December 5-8). A novel and effective algorithm for numerical optimization: melody search (MS). In *Proceeding of the 11th International Conference on Hybrid Intelligent Systems (HIS).* (pp. 109-114). Melaka, Malaysia.

Ashrafi, S. M., & Dariane, A. B. (2013). Performance evaluation of an improved harmony search algorithm for numerical optimization: Melody Search (MS). *Engineering Applications of Artificial Intelligence, 26*(4), 1301-1321.

Ashrafi, S. M., & Dariane, A. B. (2017). Coupled operating rules for optimal operation of multi-reservoir systems. *Water Resources Management, 31*(14), 4505-4520.

Ashrafi, S. M., & Kourabbaslou, N. E. (2015). An efficient adaptive strategy for melody search algorithm. *International Journal of Applied Metaheuristic Computing*, *6*(3), 1-37.

Ashrafi, S. M., Ashrafi, S. F., & Moazami, S. (2017). Developing self-adaptive melody search algorithm for optimal operation of multi-reservoir systems. *Journal of Hydraulic Structures*, *3*(1), 35-48.

Carmona, G., Molina, J. L., Bromley, J., Varela-Ortega, C., & García-Aróstegui, J. L. (2010). Object-Oriented Bayesian networks for participatory water management: two case studies in Spain. *Journal of Water Resources Planning and Management, 137*(4), 366-376.

da Silva, T. D., & Alves, C. D. M. A. (2016, May 22-26). WEAP and QUAL2K model integration for water quality evaluations as a result of urban expansion scenarios in the federal district of Brazil. In *Proceeding of the World Environmental and Water Resources Congress.* (pp. 330-338). West Palm Beach, Florida.

Dezab, C. E. (2011). *Systemic studies of Dez and Karun watersheds, demands and consumption report*. Ahvaz, Iran: Dezab Company Press.

Ehteshami, M., Biglarijoo, N., & Salari, M. (2014). Assessment and quality classification of water in Karun, Dez and Karkheh Rivers. *Journal of River Engineering, 2*(8), 23-30.

Emamgholizadeh, S., Kashi, H., Marofpoor, I., & Zalaghi, E. (2014). Prediction of water quality parameters of Karoon River (Iran) by artificial intelligence-based models. *International Journal of Environmental Science and Technology, 11*(3), 645-656.

Estalaki, S. M., Abed-Elmdoust, A., & Kerachian, R. (2015). Developing environmental penalty functions for river water quality management: application of evolutionary game theory. *Environmental Earth Sciences, 73*(8), 4201-4213.

Ghassemi, F., Jakeman, A. J., & Nix, H. A. (1995). *Salinisation of land and water resources: human causes, extent, management and case studies*. Wallingford, England: CABI Publishing.

Gitau, M. W., Chen, J., & Ma, Z. (2016). Water quality indices as tools for decision making and management. *Water Resources Management, 30*(8), 2591-2610.

Hashimoto, T., Stedinger, J. R., & Loucks, D. P. (1982). Reliability, resiliency, and vulnerability criteria for water resource system performance evaluation. *Water Resources Research*, *18*(1), 14-20.

Heydari, M., Othman, F., & Noori, M. (2013). A review of the environmental impact of large dams in Iran. *International Journal of Advancements Civil Structural and Environmental Engineering, IJACSE, 1*(1), 1-4.

Hosseini-Zare, N., Gholami, A., Panahpour, E., & Jafarnejadi, A. (2014). Pollution load assessment in the soil and water resources: a case study in Karun river drainage basin, southwest of Iran. *European Online Journal of Natural and Social Sciences*, *3*(3), 427-434.

Jafarzadeh, N., Rostami, S., Sepehrfar, K., & Lahijanzadeh, A. (2004). Identification of the water pollutant industries in Khuzestan Province. *Iranian Journal of Environmental Health, Science and Engineering, 1*(2), 36-42.

Karamouz, M., Mahjouri, N., & Kerachian, R. (2004). River water quality zoning: a case study of Karoon and Dez River system. *Iranian Journal of Environmental Health, Science and Engineering, 1*(2), 16-27.

Kerachian, R., & Karamouz, M. (2007). A stochastic conflict resolution model for water quality management in reservoir–river systems. *Advances in Water Resources, 30*(4), 866-882.

Mahmoud, M. I., Gupta, H. V., & Rajagopal, S. (2011). Scenario development for water resources planning and watershed management: Methodology and semi-arid region case study. *Environmental Modelling and Software, 26*(7), 873-885.

Moazami, S., Abdollahipour, A., Niri, M. Z., & Ashrafi, S. M. (2016). Hydrological assessment of daily satellite precipitation products over a basin in Iran. *Journal of Hydraulic Structures, 2*(2), 35-45.

Nikoo, M. R., Karimi, A., Kerachian, R., Poorsepahy-Samian, H., & Daneshmand, F. (2013). Rules for optimal operation of reservoir-river-groundwater systems considering water quality targets: application of M5P model. *Water Resources Management, 27*(8), 2771-2784.

Radmanesh, F., Zarei, H., & Salari, M. (2013). Water quality index and suitability of water of Gotvand Basin at District Khuzestan, Iran. *International Journal of Agronomy and Plant Production, 4*(4), 707-713.

Sun, N. Z., & Sun, A. (2015). *Model calibration and parameter estimation: For environmental and water resource systems*. New York, NY: Springer.

Torabian, A., & Shahavi, S., (2017). Studying Iranian drinking water quality guidelines compared to the authentic world standards. *Journal of Water and Wastewater Science and Engineering, 2*(2), 3-13.

Pertanika J. Sci. & Technol. 27 (4): 1901 - 1916 (2019)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Evaluation of Aerosol Optical Thickness over Malaysia Based on Multi-Source Ground and Satellite Data

Khaled Ali Ahmed Ben Youssef[1], Ahmad Makmom Abdullah[1]*, Helmi Zuhaidi Mohd Shafri[2] and Zulfa Hanan Ash'aari[1]

[1]*Department of Environmental Sciences, Faculty of Environmental Studies, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia*

[2]*Department of Civil Engineering, Faculty of Engineering, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia*

ABSTRACT

This study evaluates the spatiotemporal distribution of aerosol optical thickness (AOT) over Malaysia. The significance of aerosols in regional and global climate change assessment has become a pressing topic in recent climate discussions. Two different approaches are used in measuring AOT; satellite imagery and ground measurement approaches. However, the satellite approach is deemed the best way for monitoring the patterns and transport of aerosols largely due to its extensive spatial coverage and reliable repetitive measurements. The data in this study were obtained from a Sea-viewing Wide Field-of-view Sensor (SeaWiFS), a Multi-angle Imaging Spectroradiometer (MISR), and Moderate Resolution Imaging Spectroradiometer (MODIS) satellite sensors based on a NASA-operated Giovanni portal. Ground-based Aerosol Robotic Network (AERONET) datasets from two sites over the study area were also used. The results show that the highest AOT ground values of 1.93 and 2.00 were recorded in September 2015, at USM station and Kuching station, respectively. Throughout the 15 years of recorded data, the monthly average value of AOT reached its highest values in September, October, and November. In these months, the value of AOT went above 0.40, unlike in other months of the year. Significantly, the results indicate that Malaysian air quality can be evaluated based on AOT values, as these show the variation in optical properties of aerosol.

Keywords: AERONET, AOT, Malaysia, MISR, MODIS, SeaWiFS

ARTICLE INFO

Article history:
Received: 03 December 2018
Accepted: 30 January 2019
Published: 21 October 2019

E-mail addresses:
ben_youssef55@yahoo.com (Khaled Ali Ahmed Ben Youssef)
amakmom@env.upm.edu.my (Ahmad Makmom Abdullah)
helmi@upm.edu.my (Helmi Zuhaidi Mohd Shafri)
zulfa@upm.edu.my (Zulfa Hanan Ash'aari)
* Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

INTRODUCTION

Aerosol is mainly distributed between the ground surface and stratospheric layer of the earth's atmosphere. Another terminology used in place of aerosols is atmospheric particulate matter (PM) (Fang & Chang, 2010). The effects of particulate matter depends on its size. Therefore, the aerodynamic diameter of PM significantly influences PM on health. A PM with a diameter of more than 10 µm (PM_{10}) has less effect on the respiratory system of human beings because it hardly goes beyond the human nostrils. However, if the aerodynamic diameter of PM is between 2.5 µm and 10 µm, the PM will usually be absorbed into the human respiratory system (Pope et al., 1995; Shaadan et al., 2018).

Satake et al. (2004), in their study, indicated that the distribution of particulate matter in the Asian mainland was influenced by the transportation of man-made pollutants. Additionally, the problem of smoke and haze concentration in this area was aggravated by recurring slash-and-burn agricultural practices; however, the haze density depended on prevailing weather conditions. Thus, the Southeast Asian haze density is largely determined by local particulate generation, which has complicated regional air pollution issues (Xu, et al., 2015). For instance, in Klang Valley in Malaysia, the slash-and-burn practice has become a common practice in the past few years. Temporally, haze typically occurs in the southwest season from July to September (Radzi et al., 2004).

Despite the presence of different satellite sensors in the region, only a few aerosol-related studies based on satellite remote sensing data have been conducted in Malaysia due to the recurrent cloud cover in the country. The studies that have been conducted used devices such as an Advanced Very High Resolution Radiometer (AVHRR), MODIS, a Scanning Imaging Absorption Spectrometer for Atmospheric Chartography (SCIAMACHY), an Atmospheric Infrared Sounder (AIRS), and Medium Resolution Imaging Spectrometer (MERIS) data (Kanniah et al., 2016).

This paper aims to evaluate the aerosol optical thickness over Malaysia by analyzing and visualizing data from multiple satellite sources and ground measurement stations.

MATERIALS AND METHODS

Remote sensing techniques have attracted a lot of attention in the past few years. The installation of ground-based aerosol measuring stations (e.g., IMPROVE, AERONET, and EPA routine sites), and the launch of satellite remote sensing instruments (e.g., Multi-angle Imaging Spectroradiometer (MISR), and the Moderate Resolution Imaging Spectroradiometer (MODIS)) have improved our view and understanding of aerosols near the surface of the earth and its atmosphere (Di Girolamo et al., 2004; Kaufman et al., 1997). The aerosol characteristics retrieved from the MODIS sensor were obtained from seven (0.47-2.13 mm) out of 36 channels of the sensor (Chu et al., 2002; Jung et al., 2018).

The data produced by the MISR sensor is obtainable on a day-to-day basis with 17.6 × 17.6 km resolution. The MISR device obtains daytime data, which covers global areas, but its frequency depends on latitude. The interval of coverage fluctuates from 2 to 9 days in response to the intersection of their tracks poleward and their separation while moving toward the equator (Bruegge et al., 2004; Martonchik et al., 1998).

The Sea-WiFS sensor is used principally for the regular data generation of the world's ocean color and bio-optical properties. Moreover, SeaWiFS has eight spectral bands centered between 412 and 865 nm. The spatial resolution is approximately 4.5 km at nadir with a swath width of 1,502 km (±45.0°) at the equator (Wang et al., 2005).

The Aerosol Robotic Network (AERONET) is the most widely used surface measurement tool for quantifying the total column of aerosol optical characteristics. The relationship between the size and absorption of the aerosols allow for the determination of major aerosol types (Giles et al., 2012). Although the ground stations are limited in terms of spatial coverage (Qu et al., 2016), they are still useful for confirming aerosol products based on the type of satellite (Asmat et al., 2017).

Study Area

Malaysia is a Southeast Asian country located between 2.3167°N and 111.5500°E (Figure 1). The spatial extent of the study area is about 329,758 km^2 occupying the Malaysian Peninsula, which is bordered by the Asian landmass in the southern part, and the States of Sabah and Sarawak in the north-western coastal of Borneo Island. Malaysia's two separate regions are separated from each another by about 531 km (Semire et al., 2012).

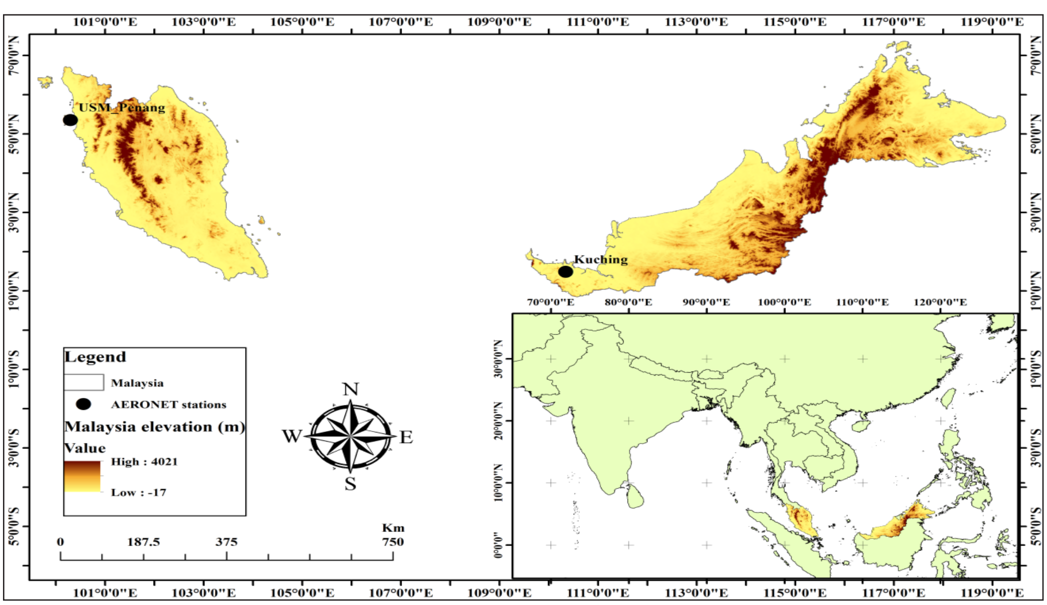


Figure 1. Study area and AERONET station location

Data and Methodology

Data Set. This study utilizes aerosol optical thickness (AOT) data from AERONET and satellites (Table 1). The data was sourced via a Giovanni online data system (Ganguly, 2016). The data was specifically retrieved from the archives of a NASA-operated Giovanni portal (http://disc.sci.gsfc.nasa.gov/giovanni). The Monthly AOT has a wavelength of 550 nm. The MODIS (Terra and Aqua satellites) (Jung et al., 2018) and SeaWiFS-derived AOT is based on a spatial resolution of 1° × 1° (Wang et al., 2005). Meanwhile, MISR is set at a spatial resolution of 0.5° × 0.5° at 555 nm (Bruegge et al., 2004). This study used different sensors to compare between different AOT values. This is essential for establishing a long-term database for atmospheric studies and will assist in improving the accuracy of results obtained using a single sensor (Prasad & Singh, 2007). From the spatial dimensions, the MISR and MODIS AOT retrieval abilities for South-East Asia was found acceptable (Xiao et al., 2009).

The data from AERONET ground-based stations are made available at three different levels; Level 1.0 (unscreened), Level 1.5 (cloud screened), and Level 2.0 (quality assured) (Holben et al., 1998). The data can be obtained from the AERONET website (http://aeronet.gsfc.nasa.gov/). AERONET is a well-established network with over 700 stations and provides standardized high quality aerosol measurements. The network is extensively used for different aerosol-related studies including satellite retrieval and validation (Cheng et al., 2012).

In Malaysia, at the moment, there are only two functional AERONET stations: Universiti Sains Malaysia (USM) Penang and Kuching stations (Figure 1). The AERONET stations use a Cimel sun photometer to measure aerosol and radiation parameters (Kanniah et al., 2016).

The most accurate way to retrieve and verify satellite remote sensing AOT data is by taking measurements using a sun photometer. The commonly used sun photometer makes a measurement every 15 min during the day in a number of wavelengths (Holben et al., 1998; Ichoku et al., 2002).

The spatial-averaged MODIS AODs are compared with the temporal-averaged sun photometer. AERONET level 2.0 AOD data are the most used data for validating the MODIS aerosol product (He et al., 2010), as it is reliable. The MODIS and SeaWiFS sensors retrieve AOT at 550-nm wavelength while MISR does so at 555 nm. This does not comply with AERONET wavelengths (348, 388, 440, 500, 675, 878, 1020, and 1648 nm). Because AERONET does not make measurements at 550 nm, the AERONET data was interpolated to 550 nm using the standard Ångström exponent, α, as defined in Equation (1).

$$\alpha = ln\ (\tau 1/\tau 2)/ln(\lambda 1/\lambda 2) \qquad (1)$$

Where, $\tau 1$ and $\tau 2$ are the AOT data at wavelengths $\lambda 1$ (500 nm) and $\lambda 2$ (675 nm), respectively. These values are the nearest available pair of bounding wavelengths from AERONET (Sayer et al., 2013).

$$\alpha = ln\,(\tau 500nm/\tau 675nm)/ln(675nm/500nm)$$

To validate the result from the MODIS sensor, the sun photometer in AERONET ground stations is used to interpolate to a common wavelength of 550 nm, where Equation (2) is applied:

$$\tau 550 \;\; = exp\,[ln(\tau 500nm) - ln(550nm/500nm)\alpha] \qquad (2)$$

The data used (Table 1) are the Version-2 direct sun algorithm Level 2 quality assured and cloud screened (Li et al., 2015). The product is available at http://aeronet.gsfc.nasa.gov/new_web/index.html

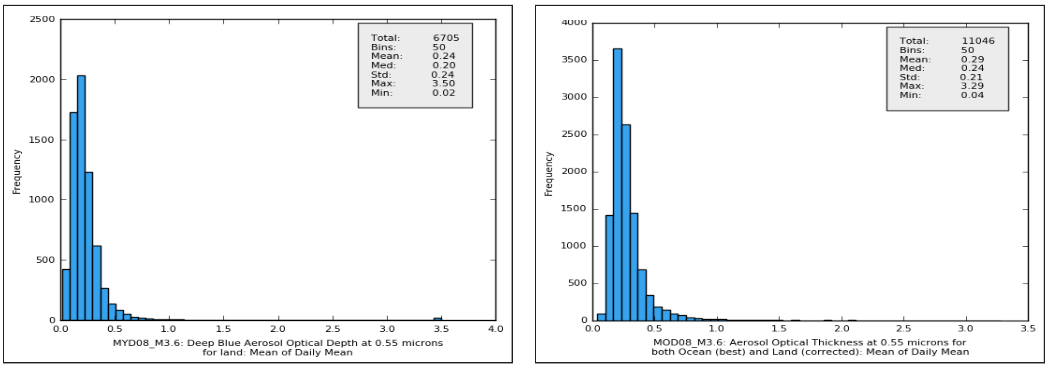
Table 1
Data sets used in the study

Satellite sensor	Special resolution	Spectral band (nm)	Data source
Satellite data			
MODIS Terra			
MODIS Aqua	1° × 1°	550	https://giovanni.gsfc.nasa.gov/giovanni/
SeaWiFS			
MISR	0.5° × 0.5°	555	
Ground data			
AERONET	Spectral band of 550 nm		http://aeronet.gsfc.nasa.gov/new_web/index.html

RESULTS AND DISCUSSIONS

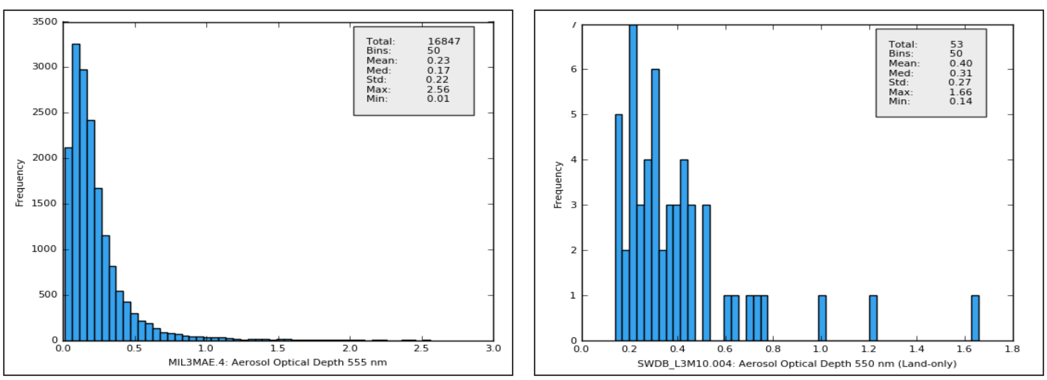
Descriptive Statistics of AOT over Malaysia Based on Satellite Remote Sensors

The spatiotemporal variabilities of aerosol optical thickness over Malaysia was assessed using information acquired from the NASA-operated GIOVANNI portal (https://giovanni.gsfc.nasa.gov/giovanni/) and AERONET ground stations. Figure 2 shows the monthly mean AOT (unitless) average results obtained from MODIS Aqua (0.24) (Figure 2 (a)), MODIS Terra (0.29) (Figure 2 (b)), and MISR (0.23) (Figure 3 (a)) from the period 2002 to 2017. The monthly mean for SeaWiFS is 0.4 from the period 1997 to 2010. The distribution of AOT skewed right—apparent in MODIS Aqua (Figure 2 (a)), MODIS Terra (Figure 2 (b)) and MISR (Figure 3 (a)). Meanwhile, for SeaWiFS, there seemed to be outliers in the AOT data to the far right, indicating values of more than 1.0 (Figure 3 (b)). The disparity in AOT values between the sensors could possibly be due to differences in platforms. This may be attributed to the large aerosol load, the complex aerosol mixtures, and the variations in climatic conditions such as precipitation, temperature, and wind pattern (Habib et al., 2018).



(a) (b)

Figure 2. Histogram of AOD at 550 nm for: (a) MODIS Aqua and (b) MODIS Terra from 2002 to 2017 (https://C.gsfc.nasa.gov/giovanni/)



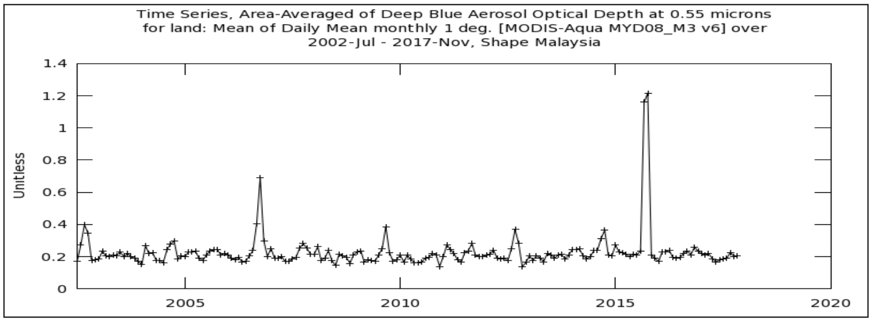
(a) (b)

Figure 3. Histogram of AOD at 550 nm for: (a) MISR at 555 nm from 2002 to 2017 and (b) SeaWiFS from 1997 to 2010 (https://giovanni.gsfc.nasa.gov/giovanni/)

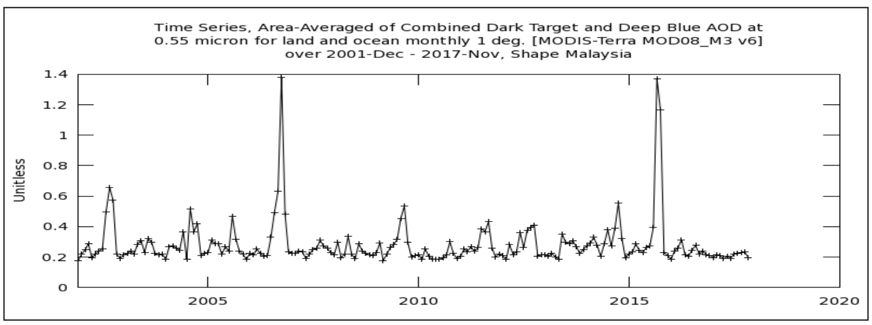
Temporal Distribution of AOT

The combined columnar aerosol disappearance is comparable to the aerosol optical thickness. This information can be harnessed to understand the spatiotemporal disparities of aerosols and their effect on radiative transfer over Malaysia (Kanniah et al., 2016).

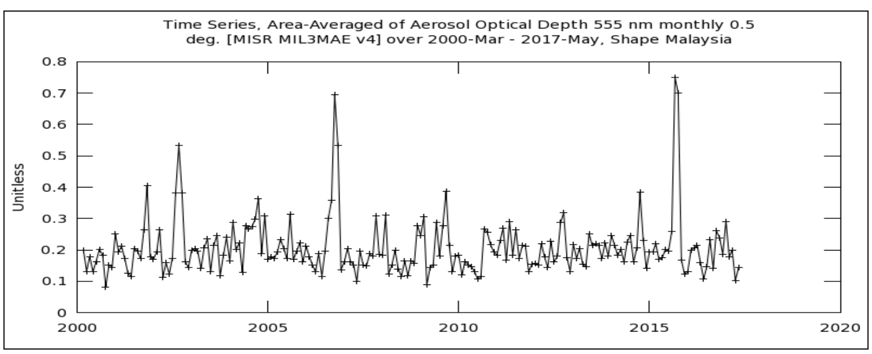
Figure 4 compares the long-term AOT temporal variability of four sensors. The monthly temporal variation of AOT for all sensors over different years was strongly similar. In general, the AOT value of the time series in all figures showed similar seasonal patterns for all sensors. Malaysia experienced two peak periods of AOT values from 2005–2007 and 2015–2016 for all the sensors. This finding is consistent with that of a previous study, which suggested that the results are due to Malaysia being exposed to haze in 2005 and 2006 (Othman et al., 2014).



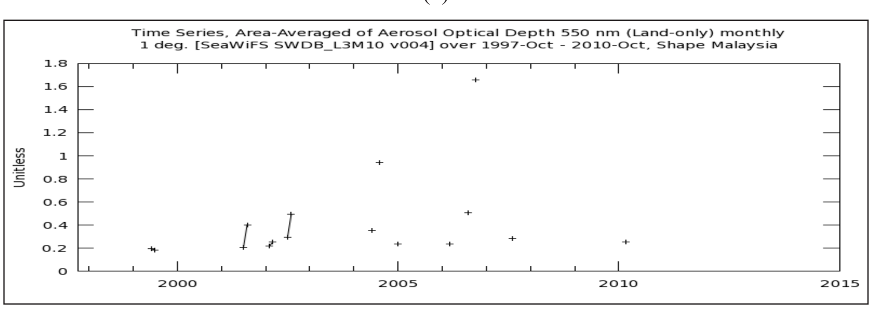
(a)



(b)



(c)



(d)

Figure 4. Monthly 1-degree time series area averaged of aerosol optical depth for: (a) MODIS Aqua Jun. 2002–Nov. 2017; (b) MODIS Terra Dec. 2001–Nov. 2017; (c) MISR Mar. 2000–May 2017; (d) and SeaWiFS from 1997 to 2010 (https://giovanni.gsfc.nasa.gov/giovanni/)

Seasonal Distribution of AOT

Recent advances in remote sensing at the time of this study has allowed for the retrieval of temporal seasonal distribution of AOT data from MODIS Terra (Figure 5 (a)), MODIS Aqua (Figure 5 (b)), MISR (Figure 5 (c)), and SeaWIFS (Figure 5 (d)). The main

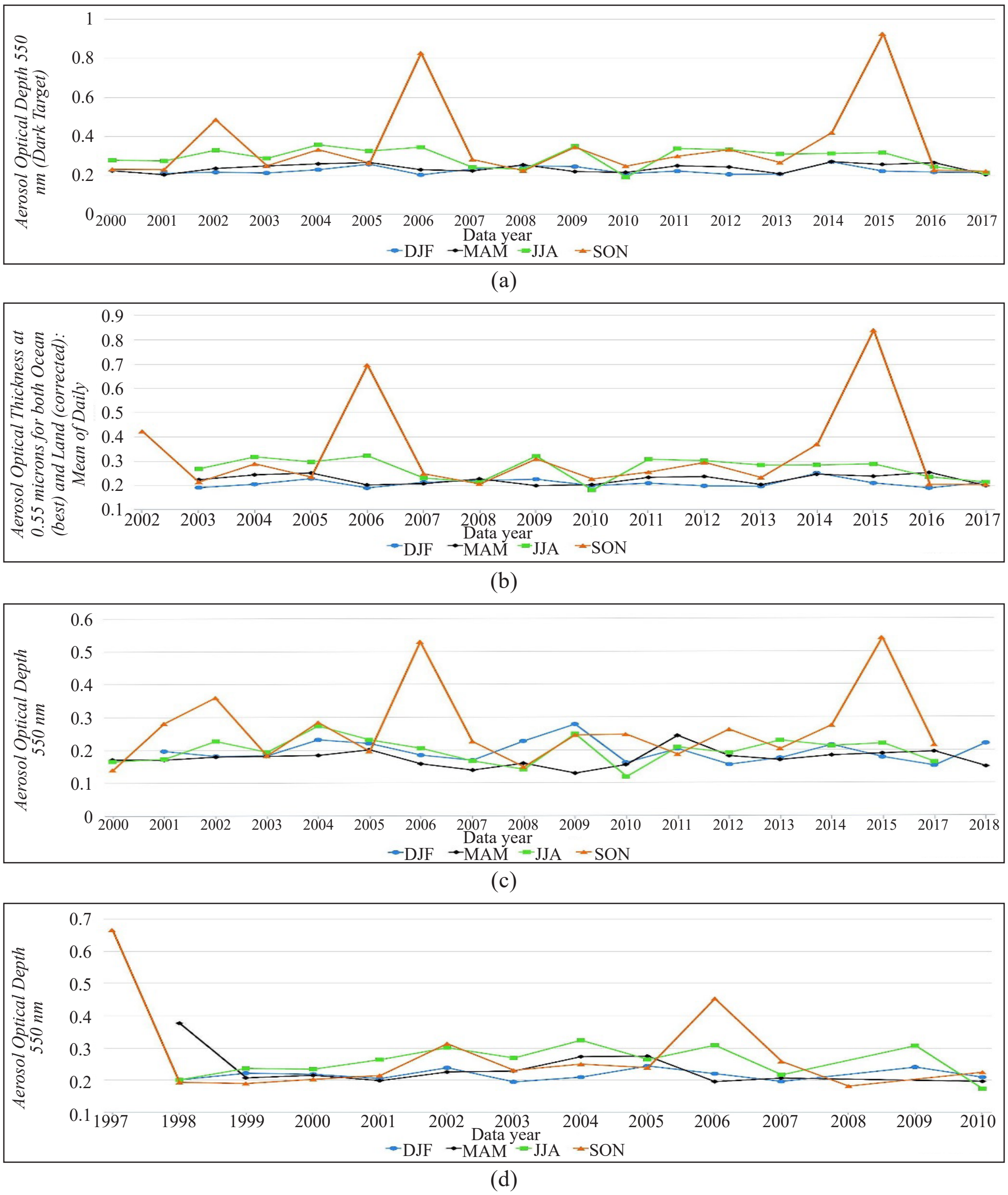


Figure 5. Seasonal time series for: (a) MODIS Terra; (b) MODIS Aqua; and (c) MISR for the period 2000–2017; as well as (d) SeaWiFS for the period 1997–2010. The months are denoted as December, January, and February (DJF), March, April, and May (MAM), June, July, and August (JJA), and September, October, and November (https://giovanni.gsfc.nasa.gov/giovanni/)

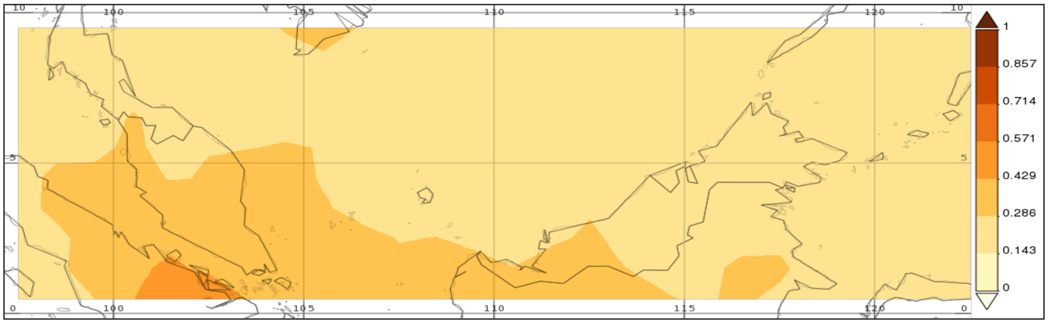
significance of these techniques is their effectiveness in detecting and comparing the averages of different seasons and AOT datasets. The results of this study show a higher value of AOT in September, October, and November (Southwest season) for all sensors. A possible explanation for this might be the regular phenomenon of Southeast Asian smog or air pollution incidents since the 1980s (Sahani et al., 2014). A substantial amount of particulate matter is released in the air from biomass burning, which moves with the help of south-westerly winds (between June and September) to Malaysia. Vehicular exhaust and burning of biomass, both locally and/or at the inter-boundary level, have consequently led to haze dispersal to Peninsula Malaysia and to Sabah and Sarawak (Sahani et al., 2014).

Tsai et al. (2011) used MODIS-derived AOD products for air quality monitoring; they found that haze layer height had larger impacts on correlation than boundary layer height, owing to an abundance of aerosols above the boundary layer. Additionally, Salinas et al. (2013) found that the temporal variability of AOD and fine-mode AOD indicated elevated levels of aerosol loading, where there were similarly high values of the Angstrom exponent number consistent with the presence of fine mode particulates. Nevertheless, these results suggest that data obtained using MODIS and SeaWiFS sensors provides more information on AOT that would be helpful in assessing the impact of haze and in predicting particulate matter in the area.

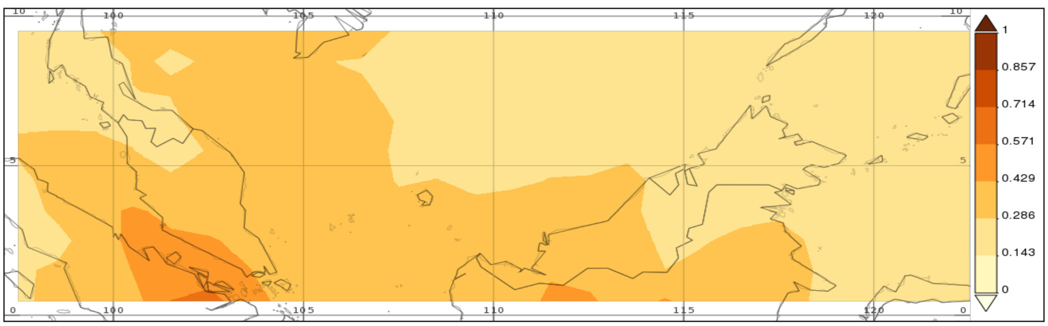
Spatial Distribution of AOT over Malaysia Area

Due to their extensive spatial coverage, satellite observations have been widely used to estimate AOT distribution. The data from the NASA-operated GIOVANNI portal (https://giovanni.gsfc.nasa.gov/giovanni/) is used to average the spatial distribution of AOT from MODIS Aqua (Figure 6 (a)), MODIS Terra (Figure 6 (b)), MISR (Figure 6 (c)), and SeaWiFS (Figure 6 (d)) over the area of Malaysia between 2005 and 2007. From these figures, it can be seen that there is slightly more satellite sensor data for Western Malaysia than its Eastern part. For example, the MODIS Terra data in Figure 6 (b) shows a clear trend of increasing AOT in Peninsular Malaysia. Kanniah et al. (2014) reported that the western sites in Peninsular Malaysia had the highest AOD 550 values from Terra MODIS, associated with densely populated, industrialized, and polluted areas. Highly populated areas have high vehicular traffic (Sahani et al., 2014). These results reflect the findings of Latif et al. (2014), who reported that urbanized areas in Peninsular Malaysia had high concentrations of aerosol.

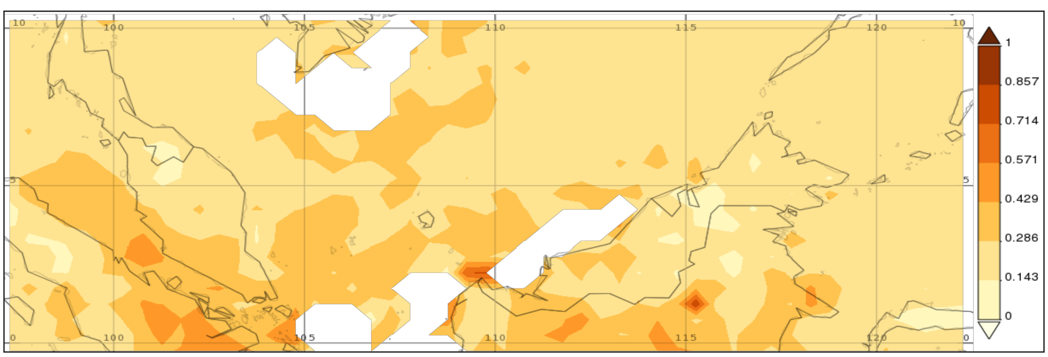
Therefore, this result indicates that satellite sensors can account for spatial variability and long-range transportation of air pollution, and thus provide coverage of local and global distribution of aerosols.



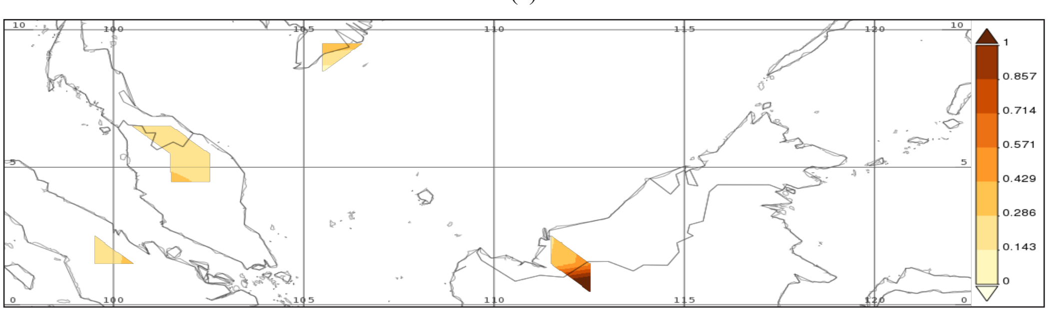
(a)



(b)



(c)



(d)

Figure 6. Spatial distribution of AOT retrieved from: (a) MODIS Aqua (550 nm); (b) MODIS Terra (550 nm); (c) MISR (555 nm); and (d) SeaWiFS (550 nm) over the area of Malaysia between 2005 and 2007 (https://giovanni.gsfc.nasa.gov/giovanni/)

Distribution of Ground-Measurement AOT

Figure 7 shows the frequency of ground measurement distribution of AOT from 2005 to 2007 for two AERONET stations in Malaysia. The AOT value in the two datasets is mostly less than 0.5, although the mean for USM station (0.66) is higher than that of Kuching station (0.27). There are similarities between the distribution of AOT in ground-based AERONET stations (Figure 7(a) & (b)) and the data distribution from satellite sensors (Figure 2 & Figure 3). The former data is skewed right, where these results support previous research that suggests that AERONET could be adopted to calibrate satellite AOT (Grosso & Paronis, 2012; Jiang et al., 2007).

The time series of AOT values from the ground from 2011 to 2018 are shown in Figure 8 and Figure 9, for the Kuching and USM stations, respectively. The AERONET data suggests a fair degree of diversity in aerosol properties in Malaysia. From the charts, it can be seen that the highest AOTs were observed in September 2015. A comparison of the findings with the seasonal time series is shown in Figures 5 (a), (b), and (c). These results are in keeping with previous studies, which suggest that the monthly concentration of major air pollutants were higher between June and September during the southwest monsoon. This is a result of air movement that transports these pollutants across the Malaysian border from Sumatra, Indonesia. Bushfires are primarily the major factor for the high density of particulate matter in the atmosphere in this area (Latif et al., 2014). However, more research on this topic needs to be done to understand the connection between satellite AOT and ground AOT measurements.

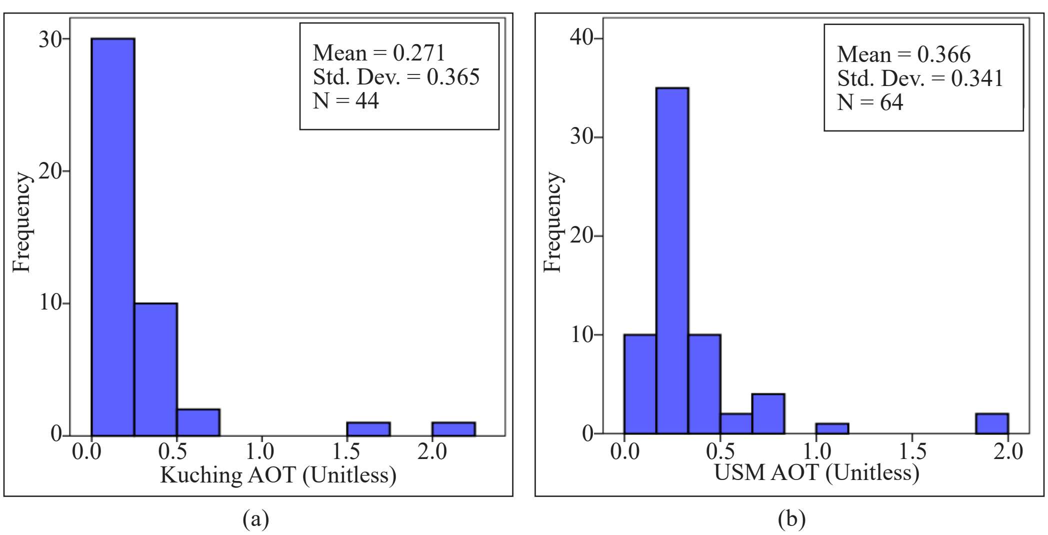


Figure 7. Histogram and description of Monthly mean AOT variability retrieved from: (a) AERONET Kuching; and (b) AERONET USM from 2011 to 2017

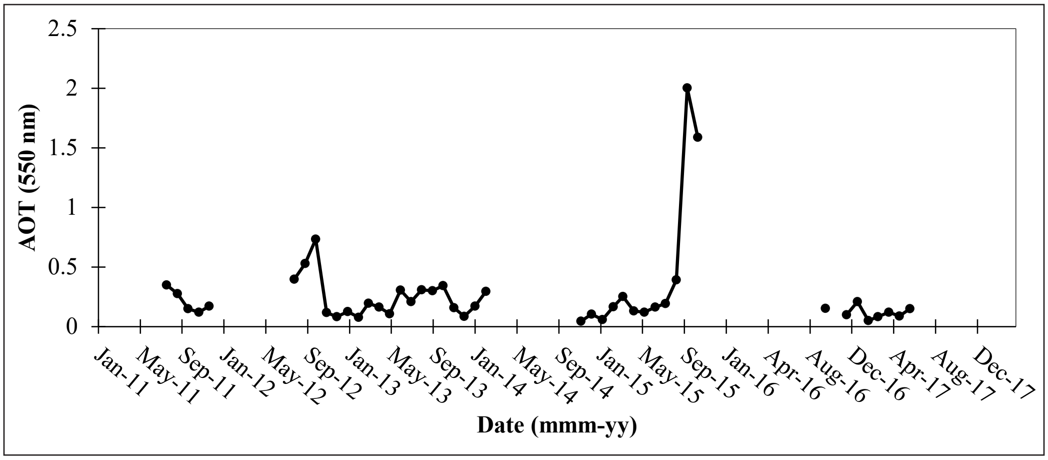


Figure 8. Monthly variation of AOD (550 nm) from 2011–2018 over KUCHING station

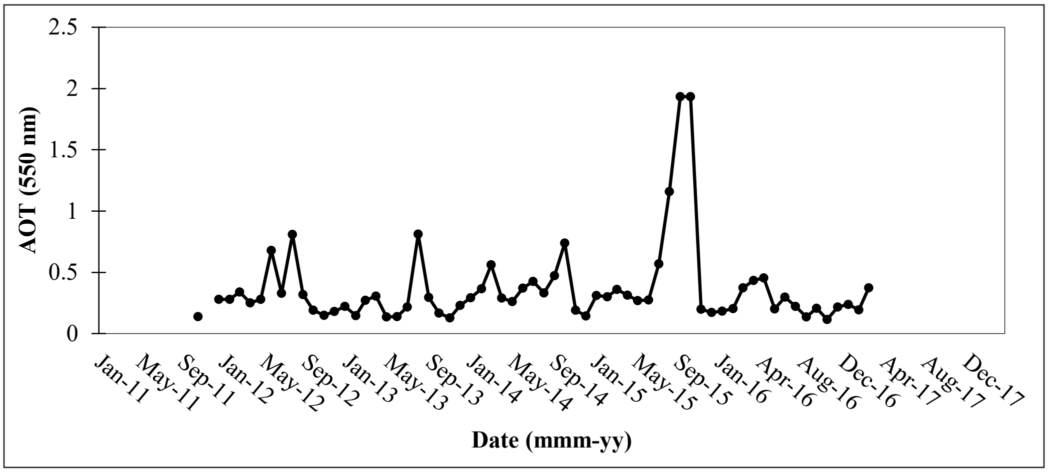


Figure 9. Monthly variation of AOD (550 nm) from 2011–2018 over USM station

CONCLUSIONS

SeaWiFS, MODIS (Terra and Aqua), and MISR aerosol measurements were used to evaluate the changeability of aerosol optical thickness over different areas of Malaysia. Both past (SeaWiFS) data and present (MODIS and MISR) data displayed an increase in AOT value between 2005 and 2007. For recent sensor data, all comparisons yielded very similar results between 2014 and 2016. The MISR data showed slightly lower AOT values than those obtained from MODIS. Nevertheless, the satellite sensor data showed very good temporal correlation, explaining the variability of AOT measured from ground AERONET data. Despite the limitation of surface observation stations, they still helped in validating the satellite data. This paper also highlighted the importance of satellite sensors in air quality monitoring in Malaysia. Future studies need to emphasize on comparing satellite and ground-based observation stations in Malaysia to better understand aerosols and their optical properties.

ACKNOWLEDGEMENTS

We thank Lim Hwee San for the establishment and maintenance of the USM Penang sites used in this study. We thank Maznorizan Mohamad and his staff for establishing and maintaining the Kuching sites used in this investigation. The analyses and visualizations used in this paper were produced with the Giovanni online data system, developed and maintained by the NASA GES DISC.

REFERENCES

Asmat, A., Jalal, K., & Ahmad, N. (2017). Aerosol radiative forcing estimation using moderate resolution imaging spectroradiometer (MODIS) in Kuching, Sarawak. *Pertanika Journal of Science and Technology, 25*(Special Issue), 125-134.

Bruegge, C., Abdou, W., Diner, D., Gaitley, B., Helmlinger, M., Kahn, R., & Martonchik, J. (2004). Validating the MISR radiometric scale for the ocean aerosol science community. *Proceedings of the International Workshop on Radiometric and Geometric Calibration* (p. 103). Mississippi, USA: CRC Press.

Cheng, T., Chen, H., Gu, X., Yu, T., Guo, J., & Guo, H. (2012). The inter-comparison of MODIS, MISR and GOCART aerosol products against AERONET data over China. *Journal of Quantitative Spectroscopy and Radiative Transfer, 113*(16), 2135-2145.

Chu, D. A., Kaufman, Y. J., Ichoku, C., Remer, L. A., Tanré, D., & Holben, B. N. (2002). Validation of MODIS aerosol optical depth retrieval over land. *Geophysical Research Letters, 29*(12), 21-24.

Di Girolamo, L., Bond, T., Bramer, D., Diner, D., Fettinger, F., Kahn, R., … Rasch, P. (2004). Analysis of Multi-angle Imaging SpectroRadiometer (MISR) aerosol optical depths over greater India during winter 2001–2004. *Geophysical Research Letters, 31*(23). http://doi: 10.1029/2004GL021273

Fang, G.-C., & Chang, S.-C. (2010). Atmospheric particulate (PM10 and PM2. 5) mass concentration and seasonal variation study in the Taiwan area during 2000-2008. *Atmospheric Research, 98*(2-4), 368-377.

Ganguly, N. D. (2016). Atmospheric changes observed during April 2015 Nepal earthquake. *Journal of Atmospheric and Solar-Terrestrial Physics, 140*(2016), 16-22.

Giles, D. M., Holben, B. N., Eck, T. F., Sinyuk, A., Smirnov, A., Slutsker, I., … Schafer, J. (2012). An analysis of AERONET aerosol absorption properties and classifications representative of aerosol source regions. *Journal of Geophysical Research: Atmospheres, 117*(D17).

Grosso, N., & Paronis, D. (2012). Comparison of contrast reduction based MODIS AOT estimates with AERONET measurements. *Atmospheric Research, 116*(2012), 33-45. http://doi:10.1016/j.atmosres.2011.09.008

Habib, A., Chen, B., Khalid, B., Tan, S., Che, H., Mahmood, T., … Butt, M. T. (2018). Estimation and inter-comparison of dust aerosols based on MODIS, MISR and AERONET retrievals over Asian desert regions. *Journal of Environmental Sciences, 76*(2018), 154-166.

He, Q. S., Li, C. C., Tang, X., Li, H. L., Geng, F. H., & Wu, Y. L. (2010). Validation of MODIS derived aerosol optical depth over the Yangtze River Delta in China. *Remote Sensing of Environment, 114*(8), 1649-1661. http://doi:10.1016/j.rse.2010.02.015

Holben, B. N., Eck, T., Slutsker, I., Tanre, D., Buis, J., Setzer, A., … Nakajima, T. (1998). AERONET-A federated instrument network and data archive for aerosol characterization. *Remote Sensing of Environment, 66*(1), 1-16.

Ichoku, C., Allen Chu, D., Mattoo, S., Kaufman, Y. J., Remer, L. A., Tanré, D., … Holben, B. N. (2002). A spatio-temporal approach for global validation and analysis of MODIS aerosol products. *Geophysical Research Letters, 29*(12), 11-14.

Jiang, X., Liu, Y., Yu, B., & Jiang, M. (2007). Comparison of MISR aerosol optical thickness with AERONET measurements in Beijing metropolitan area. *Remote Sensing of Environment, 107*(1-2), 45-53. http://doi:10.1016/j.rse.2006.06.022

Jung, C. R., Hwang, B. F., & Chen, W. T. (2018). Incorporating long-term satellite-based aerosol optical depth, localized land use data, and meteorological variables to estimate ground-level PM2.5 concentrations in Taiwan from 2005 to 2015. *Environmental Pollution, 237*, 1000-1010. http://doi:10.1016/j.envpol.2017.11.016

Kanniah, K. D., Kaskaoutis, D. G., San Lim, H., Latif, M. T., Zaman, N. A. F. K., & Liew, J. (2016). Overview of atmospheric aerosol studies in Malaysia: Known and unknown. *Atmospheric Research, 182*(2016), 302-318.

Kanniah, K. D., Lim, H. Q., Kaskaoutis, D. G., & Cracknell, A. P. (2014). Investigating aerosol properties in Peninsular Malaysia via the synergy of satellite remote sensing and ground-based measurements. *Atmospheric Research, 138*(2014), 223-239. http://doi:10.1016/j.atmosres.2013.11.018

Kaufman, Y. J., Tanré, D., Remer, L. A., Vermote, E., Chu, A., & Holben, B. (1997). Operational remote sensing of tropospheric aerosol over land from EOS moderate resolution imaging spectroradiometer. *Journal of Geophysical Research: Atmospheres, 102*(D14), 17051-17067.

Latif, M. T., Dominick, D., Ahamad, F., Khan, M. F., Juneng, L., Hamzah, F. M., & Nadzir, M. S. M. (2014). Long term assessment of air quality from a background station on the Malaysian Peninsula. *Science of the Total Environment, 482*(2014), 336-348.

Li, J., Carlson, B. E., & Lacis, A. A. (2015). How well do satellite AOD observations represent the spatial and temporal variability of PM2.5 concentration for the United States? *Atmospheric Environment, 102*, 260-273. http://doi:10.1016/j.atmosenv.2014.12.010

Martonchik, J. V., Diner, D. J., Kahn, R. A., Ackerman, T. P., Verstraete, M. M., Pinty, B., and Gordon, H. R. (1998). Techniques for the retrieval of aerosol properties over land and ocean using multiangle imaging. *IEEE Transactions on Geoscience and Remote Sensing, 36*(4), 1212-1227.

Othman, J., Sahani, M., Mahmud, M., & Ahmad, M. K. S. (2014). Transboundary smoke haze pollution in Malaysia: Inpatient health impacts and economic valuation. *Environmental Pollution, 189*, 194-201.

Pope, C. A., Thun, M. J., Namboodiri, M. M., Dockery, D. W., Evans, J. S., … Heath, C. W. (1995). Particulate air pollution as a predictor of mortality in a prospective study of US adults. *American Journal of Respiratory and Critical Care Medicine, 151*(3), 669-674.

Prasad, A. K., & Singh, R. P. (2007). Comparison of MISR-MODIS aerosol optical depth over the Indo-Gangetic basin during the winter and summer seasons (2000-2005). *Remote Sensing of Environment, 107*(1-2), 109-119.

Qu, W., Wang, J., Zhang, X., Sheng, L., & Wang, W. (2016). Opposite seasonality of the aerosol optical depth and the surface particulate matter concentration over the north China Plain. *Atmospheric Environment, 127*(2016), 90-99. http://doi:10.1016/j.atmosenv.2015.11.061

Radzi Bin Abas, M., Oros, D. R., & Simoneit, B. R. T. (2004). Biomass burning as the main source of organic aerosol particulate matter in Malaysia during haze episodes. *Chemosphere, 55*(8), 1089-1095. http://doi:10.1016/j.chemosphere.2004.02.002

Sahani, M., Zainon, N. A., Wan Mahiyuddin, W. R., Latif, M. T., Hod, R., Khan, M. F., … Chan, C.-C. (2014). A case-crossover analysis of forest fire haze events and mortality in Malaysia. *Atmospheric Environment, 96*(2014), 257-265. http://doi:10.1016/j.atmosenv.2014.07.043

Salinas, S. V., Chew, B. N., Miettinen, J., Campbell, J. R., Welton, E. J., Reid, J. S., Liya, E. Y., & Liew, S. C. (2013). Physical and optical characteristics of the October 2010 haze event over Singapore: A photometric and lidar analysis. *Atmospheric Research, 122*(2013), 555-570.

Satake, S., Uno, I., Takemura, T., Carmichael, G. R., Tang, Y., Streets, D., … Han, J. S. (2004). Characteristics of Asian aerosol transport simulated with a regional-scale chemical transport model during the ACE-Asia observation. *Journal of Geophysical Research: Atmospheres, 109*(D19), 2-17.

Sayer, A., Hsu, N., Bettenhausen, C., & Jeong, M. J. (2013). Validation and uncertainty estimates for MODIS Collection 6 “Deep Blue” aerosol data. *Journal of Geophysical Research: Atmospheres, 118*(14), 7864-7872.

Semire, F. A., Mohd-Mokhtar, R., Ismail, W., Mohamad, N., & Mandeep, J. (2012). Ground validation of space-borne satellite rainfall products in Malaysia. *Advances in Space Research, 50*(9), 1241-1249.

Shaadan, N., Deni, S., & Jemain, A. (2018). Investigating the randomness and duration of PM10 pollution using functional data analysis. *Pertanika Journal of Science and Technology, 26*(1), 299-308.

Tsai, T. C., Jeng, Y. J., Chu, D. A., Chen, J. P., & Chang, S. C. (2011). Analysis of the relationship between MODIS aerosol optical depth and particulate matter from 2006 to 2008. *Atmospheric Environment, 45*(27), 4777-4788. http://doi:10.1016/j.atmosenv.2009.10.006

Wang, M., Knobelspiesse, K. D., & McClain, C. R. (2005). Study of the Sea-Viewing Wide Field-of-View Sensor (SeaWiFS) aerosol optical property data over ocean in combination with the ocean color products. *Journal of Geophysical Research: Atmospheres, 110*(D10), 1-14.

Xiao, N., Shi, T., Calder, C. A., Munroe, D. K., Berrett, C., Wolfinbarger, S., & Li, D. (2009). Spatial characteristics of the difference between MISR and MODIS aerosol optical depth retrievals over mainland Southeast Asia. *Remote Sensing of Environment, 113*(1), 1-9.

Xu, J., Tai, X., Betha, R., He, J., & Balasubramanian, R. (2015). Comparison of physical and chemical properties of ambient aerosols during the 2009 haze and non-haze periods in Southeast Asia. *Environmental Geochemistry and Health, 37*(5), 831-841.

Pertanika J. Sci. & Technol. 27 (4): 1917 - 1927 (2019)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Growth Performance of Blood Cockle (*Tegillarca granosa*) within Kongkong Laut Estuaries, Masai, Johor

Amirul Azuan Md Joni[1], Ferdaus Mohamat Yusuff[1]*, Khairul Nizam Mohamed[1], Faradiella Mohd Kusin[1] and Syaizwan Zahmir Zulkifli[2]

[1]*Department of Environmental Sciences, Faculty of Environmental Studies, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia*
[2]*Department of Biology, Faculty of Science, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia*

ABSTRACT

A non-coastal cockle farming area such as an estuaries zone might become an alternative for continuous and sustainable cockle supply in the future. The main objective of this research is to determine the growth and mortality rate of *Tegillarca granosa* (*T. granosa)* within an estuary area. Three cockle plots were allocated along the estuary area of Kongkong Laut (P1, P2, P3) based on the geographical area, from upper to lower part of estuaries. Cockle monitoring activity was conducted from August to December 2015 for both in-situ water parameters and the cockles' growth increments. This study shows that within a brackish estuarine environment, the highest cockle growth increment was recorded within the highest water salinity trend area (26.92 ± 4.79 ppt; P2), with a shell increment of 2.70 ± 0.52 mm per month, while the lowest cockle growth increment was recorded within the lowest water salinity trend area (17.65 ± 5.73 ppt; P1) with the shell increment of 2.05 ± 0.86 mm per month. One-way ANOVA shows that there was significant difference ($p<0.05$) in growth increments among all sites of cockle plots, with significant positive correlation between the salinity level and the cockle's growth rate ($p<0.05$, $r=0.65$). This indicates that salinity level within a brackish environment plays an important role towards cockle's growth rate. Additionally, high turbidity level for a prolonged period leads to lower cockle survival rates within an area. However, further studies need to be done to look for other environmental factors that may affect cockle growth rate within an estuarine environment.

Keywords: Blood cockle, estuary, growth performance, Kongkong Laut, *Tegillarca granosa*

ARTICLE INFO

Article history:
Received: 03 December 2018
Accepted: 30 January 2019
Published: 21 October 2019

E-mail addresses:
nauza9176@gmail.com (Amirul Azuan Md Joni)
ferdius@upm.edu.my (Ferdaus Mohamat Yusuff)
k_nizam@upm.edu.my (Khairul Nizam Mohamed)
faradiella@upm.edu.my (Faradiella Mohd Kusin)
syaizwan@upm.edu.my (Syaizwan Zahmir Zulkifli)
* Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

INTRODUCTION

The blood cockle *Tegillarca granosa* (formerly known as *Anadara granosa*) is one of the major aquaculture species in Malaysia, dominating 93% of the total shellfish species production (Department of Fisheries Malaysia [DOFM], 2013). Within the west coast region of Peninsular Malaysia, blood cockle aquaculture flourishes due to the large number of juvenile cockles naturally developed in the tidal flats of this region, and many of them are collected for aquaculture. These juveniles have been supplied as seeds for aquaculture grounds within the west coast zone (Yurimoto et al., 2014a). However, since 2010, this particular sector has been significantly impacted by multiple stressors driven by the combined influence of hydrometeorological and land use change. Annual reports from the Department of Fisheries Malaysia notes that while *T. granosa* production in Malaysia has been erratic over the long term, it has dropped significantly in the past decade, specifically from 2010 onwards. At the same time, despite strategic expansion of production, productivity within licensed aquaculture plots along coastal areas has also dramatically declined (Yurimoto et al., 2014b).

Several research bodies have conducted a small number of studies aiming to have a deeper understanding regarding this issue (Harith et al., 2016; Pahri et al., 2016; Ramli, 2005; Ramli et al., 2013; Yurimoto et al., 2014a; Yurimoto et al., 2014b). Although initial findings were inconclusive, the pattern suggests that decline in *T. granosa* production may be attributed to the combined effects of a large number of environmental stressors, including pollution of coastal waters; degradation and erosion of mudflats and deterioration of shallow coastal habitats due to extreme weather events (Izura & Hooi, 2008); overstocking of *T. granosa* seed (Yurimoto et al., 2014b); high ammonia concentrations in the aquaculture systems; and rapid changes in sea surface temperature (Fadzil et al., 2010).

Therefore, since most of the major cockle's aquaculture activities in Malaysia are located within coastal area, a non-coastal zone such as estuary area may become an alternative solution for continuous and sustainable cockle supply in the future. However, as noted above, studies on the effectiveness of cockle aquaculture within a Malaysian estuarine environment remain inadequate and not well understood, particularly in terms of growth and mortality. Despite the fact that it is such valuable information, little work has been done to collect and standardize data in these two major factors as it may be utilised to predict yield under varying conditions within an estuarine environment. As part of programme to provide additional data on cockle's aquaculture within an estuary ecology, Kongkong Laut, Johor was selected as a study site, since natural cockles are present within this area but at very low abundance. Therefore, this study aims to determine the growth and mortality rate of blood cockle *T.granosa* within Kongkong Laut estuarine environment.

MATERIALS AND METHODS

Cockle Plot Set Up and Growth Measurement

Cockle plots were set up in three different locations along Kongkong Laut before starting with cockle growth measurement. These plots were labeled P1, P2 and P3 (Figure 1). This study was conducted from August 2015 until December 2015. Each plot was constructed of 4 mangrove wood pieces of 5 m height and 3 cm width as standing anchors at each corner of the plot. The mangrove wood was embedded 3 m below from the surface sediment before being enclosed with the mesh fence with an area of 1 m^2. The cockles were measured (Table 1) before being translocated from the commercial cockle farming area and about 60 cockles were dispersed into each of the plot within the Kongkong Laut estuary. Cockle growth and survival rates were collected for a 4-month period (August - December 2015). Naturally, bivalve's anterior and posterior part of the shell extend outwards as it grows, (Ubukata, 2003). Cockle growth rate was determined by measuring the shell length from the anterior to the posterior part of the shell by using the Von Bertalanffy Growth (VBG)

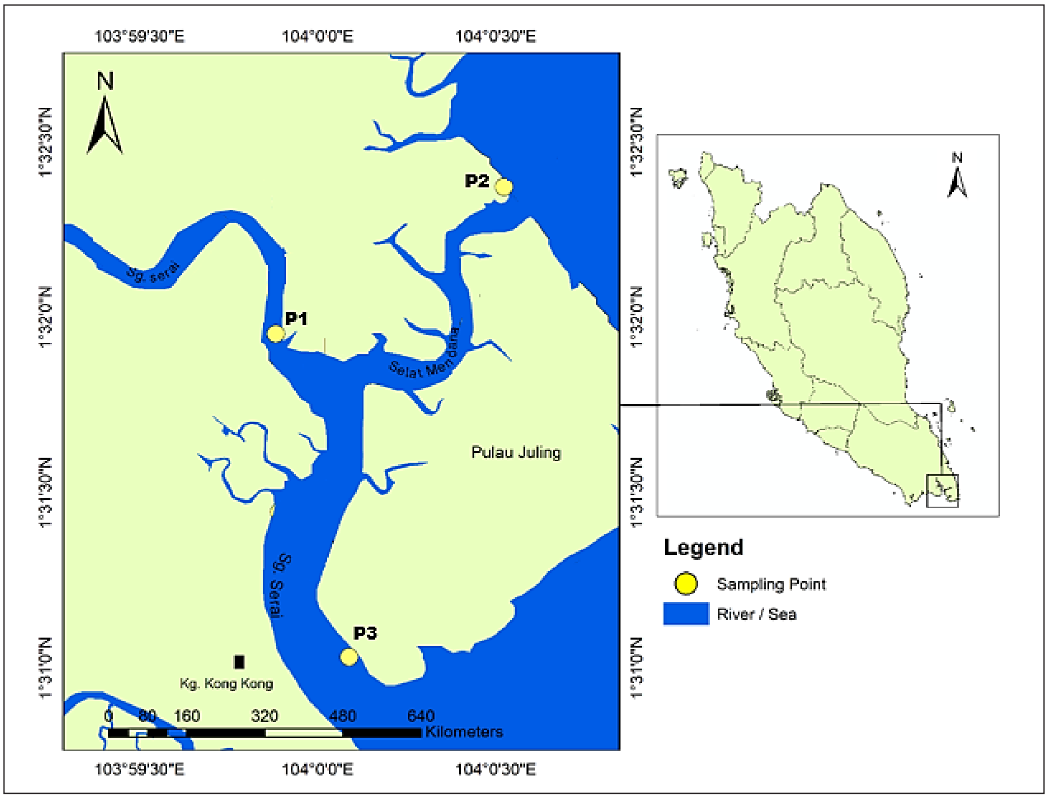


Figure 1. Location of all the cockle plots within Kongkong Laut estuary area (P1-P3)

model, as proposed by Pathansali (1966) and Broom (1982b). The VBG model has also been widely used in other fisheries related studies (De Graf & Prein, 2005). The equation is as follows:

$$\mathbf{I_t = L_i\ (1\text{-}\ e^{-Kt})} \tag{1}$$

where;

$\mathbf{I_t}$ is length at time t (Days); $\mathbf{L_i}$ is the asymptotic length (mm); and **K** is a constant indicating the rate at which the maximum size is approached.

Table 1
Total and average length (mm) of T.granosa in each of the study plot

Plot	Total Length (mm)	Average Length (mm)
P1	18.32 - 24.51	20.08
P2	17.20 – 25.22	20.29
P3	18.50 – 23.51	20.89

Water Sampling and Physicochemical Measurement

Seven water parameters were measured in this study: temperature, pH, salinity, conductivity, dissolved oxygen (DO), turbidity and water current. Unlike the data collection for cockle's growth increment, water parameter measurements were taken earlier every month within all cockle plot area (P1, P2, P3) for 8 month period (April - December 2015) in order to elucidate the pattern of water parameter (before the allocation of cockle within the plot) and clarify the factors that may influence the cockles' growth within an area (after the allocation of cockle within the plot). pH and DO were measured by using Mettler Toledo Model Sevengo Pro-SG78 (Mettler Toledo International. USA), while salinity, conductivity, temperature, turbidity and water current were measured using a Mettler Toledo Model Sevengo Pro-SG7 (Mettler Toledo International, USA) probe meter.

Statistical Analysis

Statistical analysis was performed by using IBM SPSS version 22 (SPSS Inc., Chicago, IL, USA.). One-way ANOVA was applied in order to identify the significant difference that existed within the sediment and water dataset at each sampling location. ANOVA was also used to determine the degree of similarity of the mean on the data that are continuous or normally distributed and with homogenous variance (Sekabira et al., 2010). Pearson's correlation coefficient was used in this study to determine the degree and the strength of a linear relationship between the parameters measured.

RESULTS AND DISCUSSIONS

Descriptive Statistics on Water Physicochemical Parameter

A statistical summary (mean, standard error of the mean, and standard deviation) of the selected parameters for the water samples is presented in Table 2. A total of six physicochemical variables were analysed from 10 sampling points along Kongkong Laut estuary area. Water temperature varied from 29.84 °C in P1 sampling station to 30.04 °C in P3 sampling station, which is within the portable range of 25-32°C by the World Health Organization. The pH value is within the acceptable limit of 6.5-8.5, varying between 6.79 and 7.22, with the maximum limit of 7.22 at P3 sampling station. Generally, pH value is one of the most important parameters, as it affects chemical and biological processes and temperature affects the availability of oxygen concentration in the water (Chen et al., 2006).

The concentrations of conductivity ranged from 25.58 to 43.24 ms (millisiemens). Salinity varied from 16.12 to 22.48 ppt and turbidity ranged from 16.21 to 23.97 NTU. The concentrations of these parameters were above the threshold limits of the WHO. The higher concentrations of conductivity, salinity, and turbidity indicate that these parameters were from a common source of origin (Onojake et al., 2011) and might be due to a high level of dissolved ions in the Kongkong Laut estuaries due to dissolved salt.

Table 2
Descriptive statistics; Mean and standard deviation (S.D) of water physicochemical parameters within cockle plots in 8 months period in Kongkong Laut sampling area (n=168)

Site	P1	P2	P3
Temperature (^{o}C)	29.84 ± 1.22	29.38 ± 0.52	30.04 ± 0.74
pH	6.79 ± 0.16	7.06 ± 0.34	7.22 ± 0.23
Salinity (ppt)	16.12 ± 3.90	22.33 ± 2.93	22.48 ± 2.53
Conductivity (ms)	25.58 ± 6.81	42.60 ± 3.66	43.24 ± 3.06
Dissolved Oxygen (mg/l)	6.64 ± 0.80	4.63 ± 1.97	5.11 ± 1.86
Turbidity (NTU)	21.66 ± 3.35	16.21 ± 5.98	23.97 ± 6.75
Water current (m/s)*	0.092 ± 0.001	0.199 ± 0.001	0.124 ± 0.001

Note: *3 decimal places

Physicochemical Parameter Relationship

The Pearson's correlation matrix is presented in Table 3. The correlation between the physiochemical parameters under study showed a significant positive relationship between conductivity and salinity ($r=0.894$, $p<0.01$). Conductivity is the ability of water to conduct an electric current, and the dissolved ions are the conductors. In this case, the concentration of sodium ion that derived from the brackish water might become the main contributor for the significant correlation.

There was a moderate negative relationship between DO with salinity (r= -5.02, p<0.01), conductivity (r= -5.04, p<0.01) and temperature (r= -5.07, p<0.01). This result was expected, as solubility of oxygen decreases in both warmer and saline water due to the maximum limit of air saturation within these two circumstances (Langland & Cronin, 2003; Wetzel, 2001; Yin et al., 2004). Since there was significant positive relationship between conductivity and salinity as mention earlier, a negative relationship between DO and conductivity is considered to be inevitable.

Table 3
Correlation matrix of all water physicochemical parameters within the all of the cockle plot area

	Turbidity	Conductivity	pH	Temperature	DO	Salinity	Current
Turbidity	1						
Conductivity	0.280	1					
pH	-0.051	0.160	1				
Temperature	0.457	0.524	-0.140	1			
DO	-0.470	-0.504	0.271	-0.338	1		
Salinity	0.402	0.894	0.114	0.432	-0.502	1	
Current	0.267	0.321	0.257	-0.049	-0.401	0.448	1

Cockle Growth Performance within Plot Study

The average increment of cockle growth in Plot P1, P2 and P3 is 2.05 ± 0.16, 2.70 ± 0.32 and 2.09 ± 0.27 mm per month, respectively (August - December), as can be seen in Figure 2. Among all three plots, plot P2 shows the highest average growth increment, followed by plot P3 and P1 as the lowest average growth increment. One-way ANOVA reveals that there was a significant difference (p<0.05) in growth increment of cockle between plot P2 with two other cockle plots (P1 and P3), indicating that the growth increment within the P2 plot is significantly higher compared to P1 and P3 plot. However, there was no

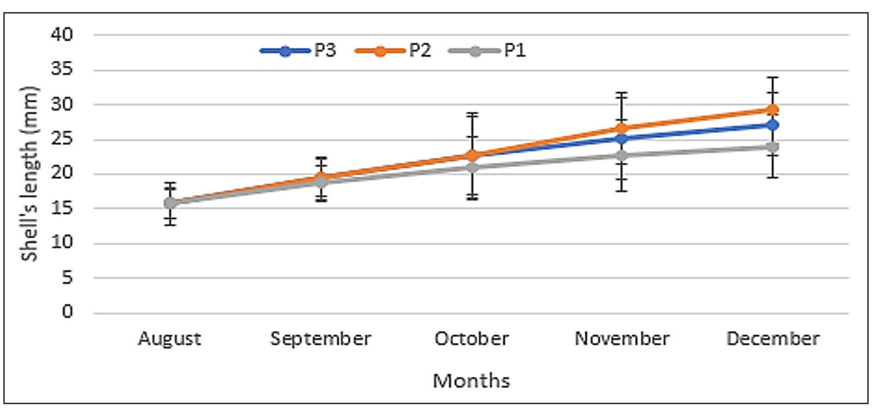


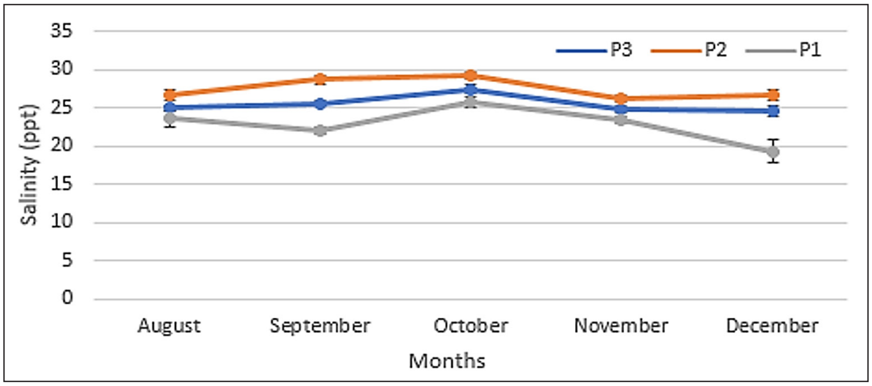
Figure 2. Growth curve of *Tegillarca granosa* within each study plot from August to December 2015, based on Von Bertalanffy growth equation

statistically certain pattern on most of the water parameters that was recorded to explain the variations of the cockle's growth, except for salinity and turbidity. Pearson's correlation analysis revealed a moderately positive correlation between the salinity level and growth rate ($p<0.05$, $r=0.65$), suggesting that salinity might be the main reason for higher cockle's growth increment within Plot P2. This result is in agreement with a study by Pathansali (1963). In his study, it was reported that the cockles' feeding rate was reduced within a lower salinity condition within a laboratory experiment, causing a slower cockle growth rate.

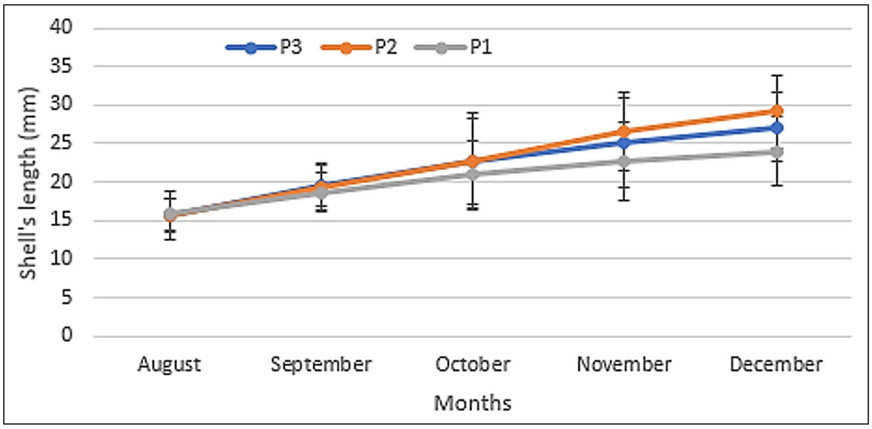
Generally, the blood cockle *T.granosa* is usually found within a coastal area or estuary area with a range of 26 to 31 ppt (Broom, 1980, 1982; Pathansali, 1963; Yurimoto et al., 2014a). In the Kongkong Laut estuary area, the water salinity ranged from 17 to 30 ppt. Among all of the cockle plots, P1 shows the lowest water salinity trend (17.65 ± 5.73 ppt), followed by P3 (23.21 ± 5.46 ppt), and P2 (26.92 ± 4.79 ppt) as the highest salinity condition within the 4-month period. Water salinity within an estuary area is not always constant, as it is also exposed directly to river and land water. During the rainy season, water salinity in some bivalve settlement area may drop as low as 10 ppt during low water of neap tides or as low as 15 ppt at high water tides (Broom, 1980; Crain et al., 2004; Day et al., 2012; Davenport & Wong, 1986). During these low salinity environments, *T. granosa* tends to close their shell valves in order to sustain the osmotic gradient between its tissue and the external medium, thus halting the feeding process until a favourable condition emerge (Davenport & Wong, 1986; Pathansali & Soong, 1958). This circumstance, however, is only restricted to a short period of time, as a prolonged low salinity environment will cause these blood cockles to be fatal (Broom, 1982a; 1982b; 1985). In this study, cockle plot at P1 area experiences the lowest average salinity (17.65 ± 5.73 ppt) within 4-month monitoring period due to the direct freshwater input from the upper stream of the estuary. The cockles in the P2 area experienced the highest average salinity (24.73 ± 5.63 ppt) within the monitoring period, as it is located nearer the Sungai Johor area which has higher salinity level compared to the inner part of the estuaries. Furthermore, the salinity trend in the cockle plot during the monitoring period indicates that the growth of the cockle increases as the salinity increases, as can be seen in Figure 3. Hence, this clearly shows that different water salinity does have an influence on the cockles' growth increment.

The average cockle survival rates in Plot P1, P2 and P3 decreased by about 15.5, 3.75 and 7.25% per month, respectively (August- December), as shown in Figure 4. Among all three plots, Plot P2 had the highest average survival rate (92 % per month) while P1 had the lowest average survival rate (73 % per month) from August to December 2015.

High turbidity trend during the four-month monitoring period (August – December 2015) suggested that turbidity might be the main reason that affecting cockles' survival within all of the plots (Figure 5). This is because of high levels of turbidity for a prolonged period have the potential to negatively affect cockle survival within a habitat drastically, as high turbidity within an environment can also potentially restrict cockle filtration activity,



(a)



(b)

Figure 3. Comparison between cockle growth (a) and salinity trend (b) in all cockle plot at Kongkong Laut area (P1, P2, P3)

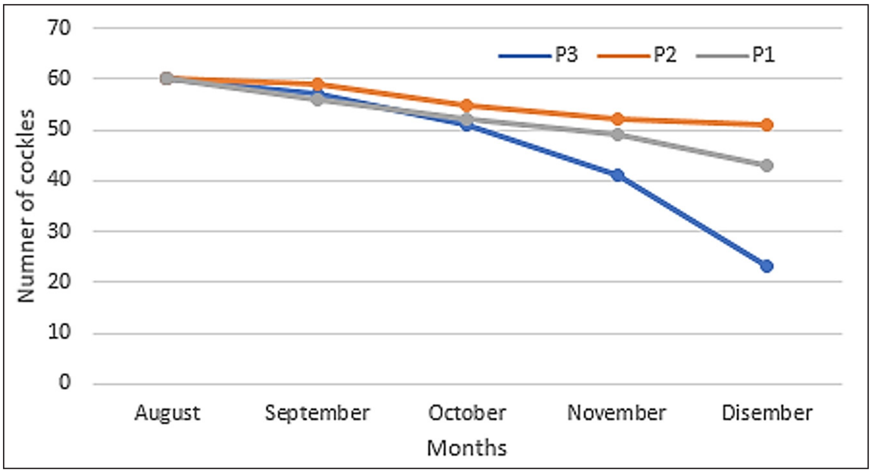


Figure 4. Survival rate comparison of *Tegillarca granosa* in plot P3, P2 and P1, based on Von Bertalanffy growth equation

thus indirectly affecting cockle feeding rate (Broom, 1980; 1985; Yurimoto et al., 2014a, 2014b). Furthermore, one-way ANOVA revealed that there was significant difference in turbidity level ($P<0.05$) between P1 and other plot study, indicating higher level or turbidity trend compared to P2 and P3. In other words, this shows that higher turbidity levels within an aquatic environment do lead to lower cockle survival within an area.

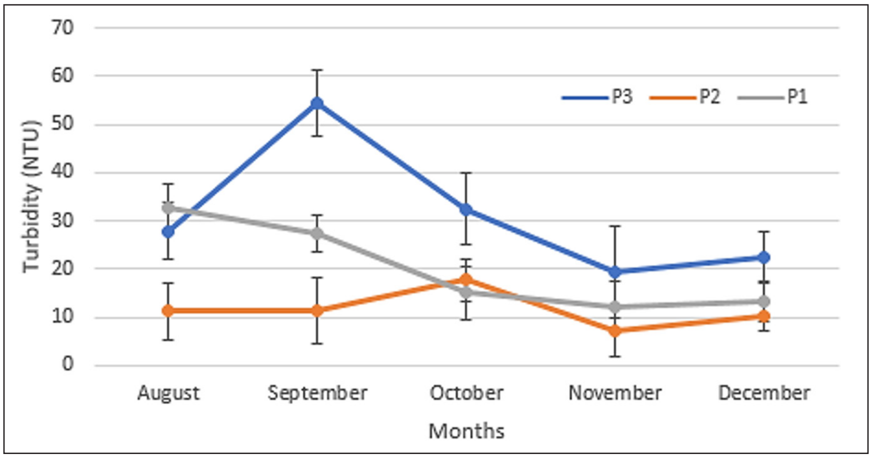


Figure 5. Turbidity level comparison within plot P1, P2 and P3 from August to December 2015

CONCLUSIONS

This study indicates that within an estuary environment, different water salinity levels have a significant influence on cockles' growth pattern, while cockle survival within an aquatic environment does influence by water turbidity level. The results suggest that lower water salinity leads to slower cockle growth increments, as it potentially affects their feeding rate negatively. On the other hand, high turbidity levels for a prolonged period lead to lower cockle survival rate within an area. Although other water parameters show no certain statistically effects on cockle growth within the study site, the range value for brackish conditions indicates that there were no extremities, and values remained within an acceptable range for cockle farming activity. A longer period of cockle growth measurement and a broader scope of environmental data are highly recommended to adapt this study in future.

ACKNOWLEDGEMENTS

The authors would like to thank the Ministry of Higher Education and Universiti Putra Malaysia for providing financial support. The project was jointly supported by FRGS 2017-1, grant no: FRGS/5540018 and Putra Graduate Initiative (IPS) grant no: IPS/9474000. Thanks, are also due to all the community members of Kong Kong Laut, Johor and Koperasi Kg. Sg. Latoh Berhad for direct and indirect support of the project.

REFERENCES

Broom, M. J. (1980). *Community and production ecology of Anadara granosa* (L.) *with particular reference to its gastropod predators* (Doctoral thesis), University of Malaya, Kuala Lumpur, Malaysia.

Broom, M. J. (1982a). Structure and seasonality in a Malaysian mudflat community. *Estuarine, Coastal and Shelf Science*, *15*(2), 135-150.

Broom, M. J. (1982b). Analysis of the growth of *Anadara granosa* (Bivalvia: Arcidae) in natural, artificially seeded and experimental populations. *Marine Ecology Progress Series. Oldendorf*, *9*(1), 69-79.

Broom, M. J. (1985). The biology and culture of marine bivalve molluscs of the genus Anadara. *ICLARM Studies and Reviews, 12*(1985)*,* 32-37.

Chen, S., Ling, J., & Blancheton, J. P. (2006). Nitrification kinetics of biofilm as affected by water quality factors. *Aquacultural Engineering*, *34*(3), 179-197.

Crain, C. M., Silliman, B. R., Bertness, S. L., & Bertness, M. D. (2004). Physical and biotic drivers of plant distribution across estuarine salinity gradients. *Ecology*, *85*(9), 2539-2549.

Day, J. W., Yáñez-Arancibia, A., Kemp, W. M., & Crump, B. C. (2012*).* Introduction to estuarine ecology. In J. W. Day, B. C. Crump, W. M. Kemp, & A. Yanez-Arencibia (Eds.), *Estuarine ecology* (pp. 1-18). Hoboke n, New Jersey: John Wiley & Sons, Inc.

Davenport, J., & Wong, T. M. (1986). Responses of the blood cockle *Anadara granosa* (L.) (Bivalvia: Arcidae) to salinity, hypoxia and aerial exposure. *Aquaculture*, *56*(2), 151-162.

Department of Fisheries Malaysia. (2013). *Annual fisheries statistics book 2013*. Department of Fisheries Malaysia. Putrajaya: Author.

De Graff, G., & Prein, M. (2005) Fitting growth with the von Bertalanffy growth function: A comparison of three approaches of multivariate analysis of fish growth in aquaculture experiments. *Aquaculture Research*, *36*(1), 100-109.

Harith, H., Husain, M. L., & Akhir, M. F. M. (2016). Coastal oceanographic processes associated with blood cockle (*Anadara granosa*) induce spawning season in Kapar, Selangor, Malaysia. *Journal of Ocean Engineering and Science*, *1*(4), 289-299.

Izura, S. N., & Hooi, T. K. (2008, July). Shaping the future of cockle industry in Malaysia. In *Proceedings of the National Fisheries Symposium (NAFIS)* (pp. 14-16). Kuala Terengganu, Terengganu.

Langland, M., & Cronin, T. (2003). *A summary report of sediment processes in Chesapeake Bay and watershed* (No. 2003-4123). Virginia, US: United States Geological Survey.

Pahri, S. D. R., Mohamed, A. F., & Samat, A. (2016). Preliminary water quality study in cockle farming area in Malaysia: A case study in Jeram, Selangor. *Aquaculture, Aquarium, Conservation and Legislation-International Journal of the Bioflux Society, 9*(2), 316-325.

Pathansali, D., & Soong, M. K. (1958). Some aspects of cockle (Anadara granosa L.) culture in Malaya. *Proceedings of the Indo-Pacific Fisheries Council*, *8*(2), 26-31.

Pathansali, D. (1963). On the effect of salinity changes on the activity of the cockle, *Anadara granosa* (L). *Malaysian Agricultural Journal, 44*(1963)*,* 18-25.

Pathansali, D. (1966). Notes on the biology of the cockle, *Anadara granosa* L. *Proceedings of the Indo-Pacific Fisheries Council*, *11*(2), 84-98.

Ramli, M. F. S. B. (2005). *Impacts of coastal land reclamation on the fisheries of Mukim Lekir, Malaysia* (Doctoral thesis), University of Hull, Cottingham, United Kingdom.

Ramli, M. F. S., Abu Hassan, F. R., & Saadon, M. N. (2013). Declining production of cockles in relation to ammonia concentrations in Sungai Buloh River, Selangor. *Journal of Environment and Earth Science*, *3*(10), 1-5.

Sekabira, K., Origa, H. O., Basamba, T. A., Mutumba, G., & Kakudidi, E. (2010). Assessment of heavy metal pollution in the urban stream sediments and its tributaries. *International Journal of Environmental Science & Technology*, *7*(3), 435-446.

Ubukata, T. (2003). A theoretical morphologic analysis of bivalve ligaments. *Paleobiology*, *29*(3), 369-380.

Wetzel, R. G. (2001). *Limnology: Lake and river ecosystems* (3rd Ed.). San Diego, USA: Academic Press.

Yassin, Z. (2016). *Environmental destruction cause of decline in cockle production*. Retrieved May 7, 2018, from https://www.consumer.org.my/index.php/ development/environment/1012-environmental-destruction-cause-of-decline-in-cockle-production

Yin, K., Lin, Z., & Ke, Z. (2004). Temporal and spatial distribution of dissolved oxygen in the Pearl River Estuary and adjacent coastal waters. *Continental Shelf Research*, *24*(16), 1935-1948.

Yurimoto, T., Kassim, F. M., Fuseya, R., & Man, A. (2014a). Mass mortality event of the blood cockle, *Anadara granosa*, in aquaculture ground along Selangor coast, Peninsular Malaysia. *International Aquatic Research*, *6*(4), 177-186.

Yurimoto, T., Kassim, F. M., Man, A., & Fuseya, R. (2014b). Spawning season and larval occurrence of blood cockle (*Anadara granosa*) off the Selangor coast, Peninsular Malaysia. *International Journal of Aquatic Biology*, *2*(6), 299-304.

Pertanika J. Sci. & Technol. 27 (4): 1929 - 1939 (2019)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Spatial and Vertical Metals Distribution in Sediment Cores from Kongkong Laut Estuary, Masai, Johor

Sara Azuana Khalidah Khalid[1], Ferdaus Mohamat Yusuff[1,2]*, Faradiella Mohd Kusin[1,2], Aziz Arshad[3] and Zulfa Hanan Asha'ari[1]

[1]*Department of Environmental Sciences, Faculty of Environmental Studies, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia*
[2]*Environmental Forensics Research Unit (ENFORCE), Faculty of Environmental Studies, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia*
[3]*Department of Aquaculture, Faculty of Agriculture, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia*

ABSTRACT

Sediment has been widely used as a means for assessing the impact of anthropogenic activities on aquatic environment, with metal pollution being the most well-known threat to aquatic environments. Therefore the aim of this study is to identify and determine metal concentrations, specifically that of copper (Cu), chromium (Cr), cadmium (Cd), iron (Fe), and zinc (Zn) on four sediment cores samples of KongKong Laut Masai Johor. The aqua regia method had been used to analyze Cu, Cr, Cd, Fe and Zn concentrations in sediment samples. Metal concentrations had then been determined through the use of flame atomic absorption spectrometry. The degree of metal contamination in sediment core samples had been determined through using Enrichment Factors (EF) and Geoaccumulation Indices (I_{geo}). From this study, the mean and standard deviation of metal concentrations in each sediment core sample were shown to be lower than the average shale concentration and lower than both the effects-range low (ERL) and effects-range median (ERM) parameters for aquatic toxicity. This information could be useful, serving as a baseline for evaluating the potential impacts of future development in the area.

Keywords: Enrichment factor, geo-accumulation index, KongKong Laut, metal, pollution, sediment core

ARTICLE INFO

Article history:
Received: 03 December 2018
Accepted: 30 January 2019
Published: 21 October 2019

E-mail addresses:
sara.azuana@gmail.com (Sara Azuana Khalidah Khalid)
ferdius@upm.edu.my (Ferdaus Mohamat Yusuff)
diellakusin@gmail.com (Faradiella Mohd Kusin)
azizarshad@upm.edu.my (Aziz Arshad)
zulfa@upm.edu.my (Zulfa Hanan Asha'ari)
* Corresponding author

INTRODUCTION

Metal contamination in sediment is a worldwide concern, due to its potential to remobilization, bioaccumulation and

ISSN: 0128-7680
e-ISSN: 2231-8526

biomagnification in aquatic ecosystems (Al-Mur et al., 2017). Metals can be divided into essential and non-essential elements, both having the potential to be toxic to living organisms when their concentration exceeds from the permissible limit, since both are persistent, which can be bioaccumulated and biomagnified through the food chain (Lim et al., 2012; Saraee et al., 2011).

The determination of metals distribution in core sediment samples can provide information about previous and present levels of contamination in aquatic environments (Al-Najjar, 2011; Cho et al., 2015). Therefore, the study of core sediment samples could provide historical evidence and historical records of metals distribution over the time (Williams & Block, 2015).

Agricultural activities such as palm oil plantations, and increasing interest in eco-tourism sectors in KongKong Laut Masai, have led to the vigorous development such as jetty, restaurants and raft houses along the KongKong Laut River, which can have a negative impact on the area's aquatic ecosystem in this area and directly expose risks to the food chain. Therefore the study of metal in the core sediment sample is needed, so as to provide useful information regarding anthropogenic activities along the KongKong Laut River.

Several studies have been conducted on surface sediment along the KongKong Laut River (Wan & Mohamat-Yusuff, 2014; Zulkifli et al., 2010). However, no study has yet been conducted on sediment cores along the KongKong Laut River. Therefore, the goal of the present study is to assess the spatial and vertical distribution of Zinc (Zn), Cadmium (Cd), Iron (Fe), Copper (Cu) and Chromium (Cr) in sediment cores, providing information which may serve as a baseline for evaluating potential impacts of future development in the study area.

MATERIALS AND METHODS

Sample Collection and Analytical Procedure

The sampling for sediment core samples was conducted in December 2016, involving four sampling stations points along the KongKong Laut River (Figure 1). Each sampling station was adjacent to a different type of human activity which it was believed could release significant amounts of anthropogenic pollution, such as metal pollution (Table 1).

Core samples were collected at a distance ranging from 1.5m to 2m from the low tide mark. The core sample lengths from each station varied due to the stratum differences existing in each station. Each core was sliced into 3cm sections. The core samples were then placed in a pre-labeled zip lock plastic bag, and kept in an icebox during transportation. At the laboratory, all samples were kept frozen at -20°C for further analysis. All reagents used were analytical-reagent grade certified. Samples were oven dried at 60°C, until a constant weight was obtained. Samples were then grained and sieved through a 63µm mesh size.

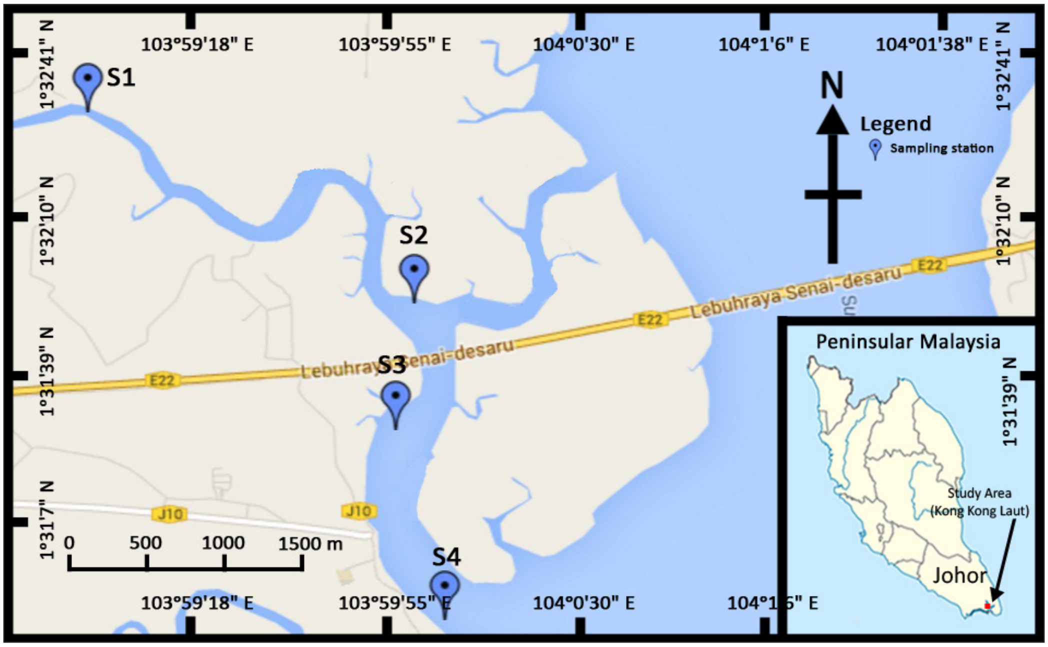


Figure 1. Sampling station for sediment core samples along the KongKong Laut River

Table 1
Coordinates and descriptions of the sampling sites

Sampling station	Abbreviation	Coordinate	Site description
Serai River	**S1**	1° 32' 31.0"N 103° 58' 57.1"E	A site near to palm oil plantation
Serai River	**S2**	1° 31' 51.1"N 103° 59' 58.4"E	A site near to waterway transportation and fishing activities
Mendana Strait	**S3**	1° 31' 28.0"N 103° 59' 59.0"E	A site near to raft house and jetties
Mendana Strait	**S4**	1° 30' 49.9"N 104° 00' 05.0"E	A site near to restaurants and abandon ship

The aqua regia method was used to determine the pseudototal trace metals in sediment samples (Wan & Mohamat-Yusuf, 2014; Zulkifli et al., 2010). The aqua regia solution (12 ml), consisting of a 3:1 ratio of hydrochloric acid to nitric acid, was added to one gram of the 63µm size sediment sample, which was then left to stand overnight. Afterwards, the sample was heated to 50°C in the digestion block for 1 hour. Then the temperature was increased to 140°C for 3 hours. The extract was filtered through a 0.45µm Whatman No.1 filter paper, and diluted in 50ml milli-Q water. All glassware and equipment were acid-soaked with 10% HNO_3 for 24 hours prior to use. The quality of the method used was checked through using the Standard Reference Material 1646a (SRM 1646a). Samples were stored at 4°C prior to analysis using an atomic absorption spectrophotometer (AAS) (Perkin-Elmer Model AAnalyst 800 series, USA). The recoveries of measured elements ranged from 99.6% to 100.54% (Table 4).

Organic matter was determined through the loss on ignition method (Al-Trabulsy et al., 2013). Particle size analysis was conducted through the pipet method (Gee & Bauder, 1986). In this study, statistical analysis was performed using the statistical package for the social sciences (SPSS) version 22, and Microsoft Office Excel 2010 (Microsoft Corporation, USA). The degree of trace metal contamination in the sediment core samples was assessed through using the Enrichment Factor (EF) (Zulkifli et al., 2010) and the Index of Geoaccumulation (I_{geo}) (Muller, 1969).

Calculation of Enrichment Factor and Geo-aAccumulation Index

$$EF = \frac{(C_x/C_{Fe})sample}{(S_x/S_{Fe})shale}$$

Where;

C_x = concentration of the measured trace element in the study area (in mg/kg)

C_{Fe} = concentration of the measured Iron in the study area (in mg/kg)

S_x = concentration of the reference trace elements in the reference environment (in mg/kg)

S_{Fe} = concentration of the Iron in the reference environment (in mg/kg)

Iron (Fe) was chosen as the reference element for this present study, because it is associated with a fine solid surface, such as the mangrove sediment found in this study area. This is due to its geochemistry being close to many trace metals, and also due to its natural concentration which tends to be uniform in natural sediment (Daskalakis & O'Connor, 1995; Naji & Ismail, 2011). The EF result can be classified into five enrichment categories, as presented in Table 2, whereas the I_{geo} result can be classified into seven categories shown in Table 3.

Table 2

The degree of metal enrichment based on EF categories

Enrichment factor (EF) value	Enrichment categories
< 2	Depletion to minimal enrichment
2-5	Moderate enrichment
5-20	Significant enrichment
20-40	Very high enrichment
>40	Extremely high enrichment

$$I_{geo} = \frac{log_2\, C_x}{1.5\,.S_x}$$

Where:

C_x = Measured trace element in the sample (in mg/kg)

S_x = Background or pristine value of the trace element (in mg/kg)

1.5 = Factor introduced in order to minimize the effect of possible variation in the background value that may be attributed to lithologic variations in the sediments (Stoffers et al., 1986)

Table 3
The degree of trace metal contamination based on I_{geo} contamination classification

I_{geo} value	**I_{geo} Class**	**Designation of sediment quality**
≤0	0	Uncontaminated
0-1	1	Uncontaminated to moderately contaminated
1-2	2	Moderately contaminated
2-3	3	Moderately to strongly contaminated
3-4	4	Strongly contaminated
4-5	5	Strongly to extremely contaminated
>5	6	Extremely contaminated

RESULTS AND DISCUSSIONS

Results of the descriptive analysis of trace element concentrations from four sediment core samples have been presented in Table 4. From the five trace elements that have been studied, the highest mean concentration was Fe (3.80 ± 0.45% dry weight), as Fe is a vital element for flora and fauna in maintaining metabolic and cellular functions. This is followed by Zn (75.07 ± 5.67 mg/kg), Cu (18.34 ± 2.07 mg/kg), Cr (13.22 ± 2.72 mg/kg), and Cd (0.21 ± 0.08 mg/kg). All mean and standard deviations of trace metal concentration were still generally lower than the average shale concentrations, as reported by Turekian and Wedepohl (1961), and were also lower than the effects range low (ERL) and effects range median (ERM) values (Long et al., 1995).

Table 4
Mean and standard deviation of trace metal concentration in sediment core samples from all locations at KongKong Laut River are expressed in units of mg/kg with the exception of Fe (%)

Site	Zn	Cd	Fe[a]	Cu	Cr
S1 (n=7)	64.85 ± 2.68	0.19 ± 0.05	4.26 ± 0.40	23.06 ± 3.19	19.73 ± 3.73
S2 (n=13)	78.41 ± 7.23	0.21 ± 0.07	3.73 ± 0.61	17.55 ± 1.10	8.99 ± 2.97
S3 (n=9)	81.69 ± 4.98	0.22 ± 0.08	3.35 ± 0.20	18.37 ± 1.80	12.86 ± 2.16
S4(n=7)	75.32 ± 7.78	0.23 ± 0.1	3.82 ± 0.52	14.38 ± 2.17	11.30 ± 2.03
Average	75.07± 5.67	0.21 ± 0.08	3.80 ± 0.43	18.34 ± 2.07	13.22± 2.72
Average shale[b]	95	0.3	4.72	45	90
ERL[c]	150	1.2	na	34	81
ERM[c]	410	9.6	na	270	370
Recovery of analysis (%)	100.22	100.54	99.6	99.74	100.43

Notes: [a] % dry weight; [b] Turekian and Wedepohl (1961); [c] Long et al., (1995); na = not available; n = number of sub-samples

The spatial distribution of metals in the upper surface samples (0-3 cm depth) from all sites is presented in Figure 2. Cu concentration is higher at S1, which may be derived from the fertilizer used on the palm oil plantation near this site. The Zn concentration is higher at S3 when compared to other stations, which might be derived from the water transportation activities conducted at this sampling station, located near to water ways and jetty operations. Cd concentration is higher at S4, which may be the result of untreated sewage effluent, as this is located near the restaurant and the raft-house. However, when compared to the average shale, ERL and ERM, all metals concentrations are still lower than the threshold limit.

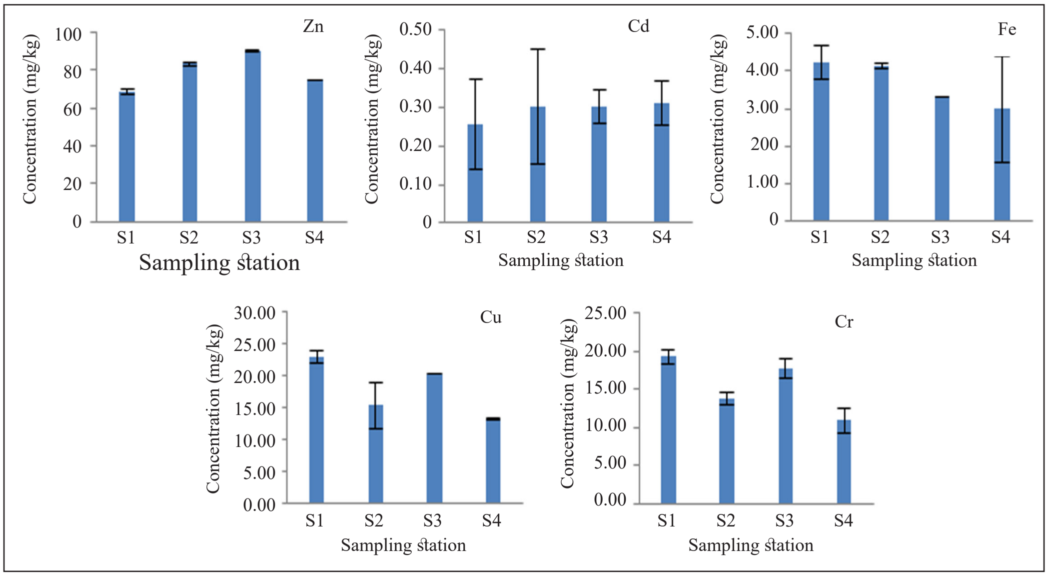


Figure 2. Metal concentration in the surface samples (0-3 cm depth) from all four sites

The vertical profiles of the Zn, Cd, Fe, Cu, and Cr concentrations from the four cores are shown to have a fluctuating pattern in regards to depth, without showing obvious increasing or decreasing trends (Figure 3). The vertical distribution of Cd in all cores showed a decreasing pattern from upper part toward the bottom part indicating an increasing enrichment of this element at all cores. This might be derived from the combustion of fuel used for water transportation and the burning of waste such as plastics, batteries, construction material and other matter by local people near the estuary, which may enter the estuary during the rainy time season. The vertical profile for Cu showed an increasing pattern of concentration downward in all cores, except for core S3. Cu is an essential nutrient and is required by organisms in small amounts (White & Rainbow, 1985). Due to the biological requirements of aquatic organisms, the concentration of Cu in surface sediment is lower than that of the bottom sediment. However, at S3 the concentration of Cu was higher in the surface sediment, which might be due to anthropogenic activities occurring in this area, such as jetty operations and heavy water transportation. The distribution pattern of Cr at all core sites showed a decreasing trend from the upper part to the lower part, indicating a reduced discharge of the Cr effluent source in the study area.

To determine the relationship between metals and sediment texture, a Pearson correlation analysis was performed. Fe shows strong positive correlations with Cr at core S1 (see Table 5) indicating that the sources of both elements come from the same sources which are believed to come from the natural sources as the average concentrations of both elements is lower than the permissible limit shown in Table 4. From the correlation analysis, there is a significant, strong positive correlation between Zn and Fe from the cores S2 (see Table 6) and S4 (see Table 8), indicating that it is scavenged by Fe oxides. This correlates with the findings of Laluraj and Nair (2006), who determined that geochemical variations in metal distribution was associated with the precipitation or dissolution of Fe oxides, acting as a source or sink for metal. Correlation analysis had shown that sand was significantly,

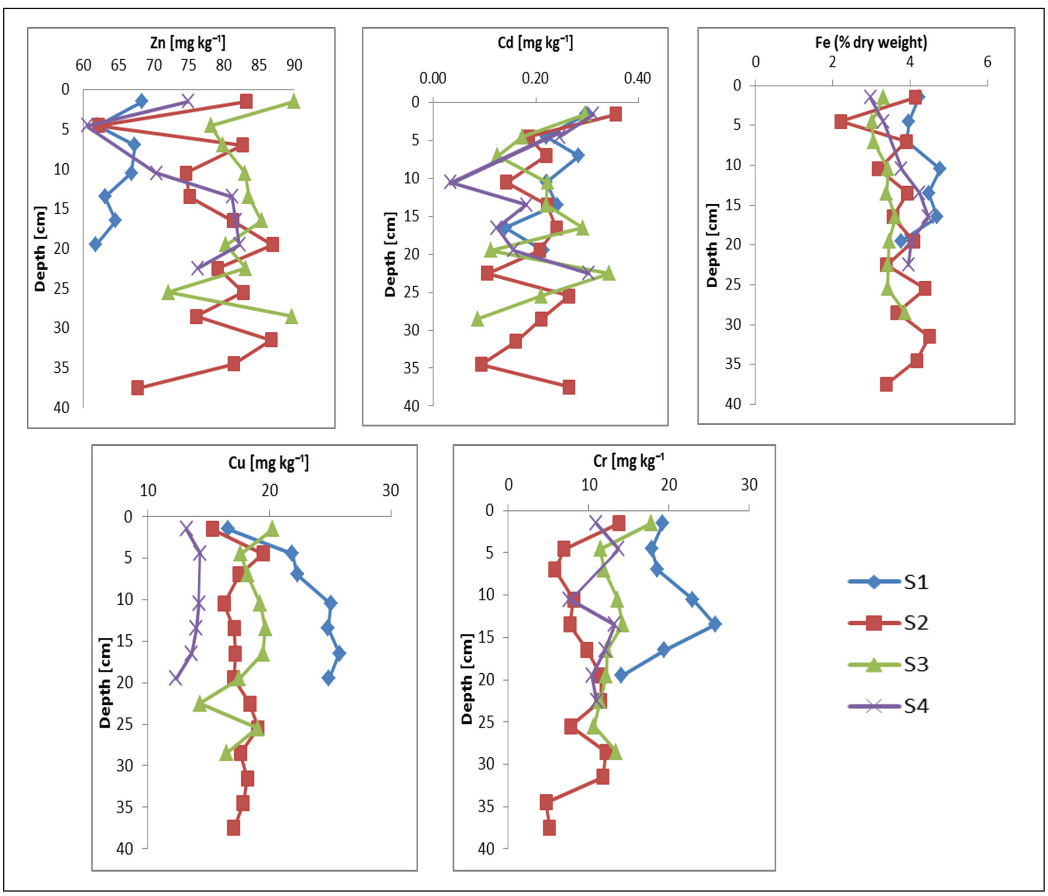


Figure 3. Depth profiles of Zn, Cd, Fe, Cu and Cr concentrations from station S1, S2, S3 and S4

negatively correlated with silt and clay in all cores, except for S2 and S3. This showed that an increased percentage of silt and clay would lead to a decreased percentage of sand. Cd and Cr had shown significant, positive correlations with Zn at core 3 (see Table 7), indicating that the source of Cd and Cr at S3 came from both anthropogenic and natural. A strong positive correlation between Fe and Cu from core S4 indicates that Cu elements are believed to be derived from a natural source.

Table 5
Pearson correlation coefficient values between metals, organic matter, clay, silt and sand for core 1

	Zn	Cd	Fe	Cu	Cr	OM	Clay	Silt	Sand
Zn	1								
Cd	0.63	1							
Fe	0.32	-0.22	1						
Cu	-0.07	-0.67	0.12	1					
Cr	0.36	0.11	0.86*	0.43	1				
OM	0.04	0.58	-0.29	-0.64	0.07	1			
Clay	0.29	0.23	0.32	-0.36	0.29	0.61	1		
Silt	0.07	-0.18	0.50	0.11	0.61	0.57	0.79*	1	
Sand	0.00	-0.07	-0.40	0.11	-0.50	-0.68	-0.86*	-0.94**	1

Notes: *significant level of $p<0.05$, **significant level of $p<0.01$

Table 6
Pearson correlation coefficient values between metals, organic matter, clay, silt and sand for core 2

	Zn	Cd	Fe	Cu	Cr	OM	Clay	Silt	Sand
Zn	1								
Cd	0.09	1							
Fe	0.81**	0.17	1						
Cu	-0.06	-0.36	0.05	1					
Cr	0.45	0.12	0.19	0.20	1				
OM	0.29	-0.00	0.35	-0.03	0.15	1			
Clay	0.37	-0.19	0.07	-0.07	0.13	0.42	1		
Silt	-0.30	0.20	-0.07	-0.08	-0.07	0.13	0.42	1	
Sand	-0.17	-0.23	0.10	0.39	-0.59*	0.50	0.23	-0.33	1

Notes: *significant level of $p<0.05$, **significant level of $p<0.01$

Table 7
Pearson correlation coefficient values between metals, organic matter, clay, silt and sand for core 3

	Zn	Cd	Fe	Cu	Cr	OM	Clay	Silt	Sand
Zn	1								
Cd	0.70*	1							
Fe	0.31	0.24	1						
Cu	0.58	0.33	-0.09	1					

Table 7 *(continue)*

	Zn	Cd	Fe	Cu	Cr	OM	Clay	Silt	Sand
Cr	0.77*	0.25	0.04	0.77*	1				
OM	-0.8	0.40	-0.51	0.13	-0.15	1			
Clay	-0.27	-0.18	0.31	-0.58	-0.23	-0.22	1		
Silt	0.27	0.18	-0.31	0.58	0.23	0.22	-1.0**	1	
Sand	-0.43	-0.32	0.02	-0.10	-0.45	0.07	0.07	-0.07	1

Notes: *significant level of $p<0.05$, **significant level of $p<0.01$

Table 8
Pearson correlation coefficient values between metals, organic matter, clay, silt and sand for core 4

	Zn	Cd	Fe	Cu	Cr	OM	Clay	Silt	Sand
Zn	1								
Cd	-0.25	1							
Fe	0.79*	-0.54	1						
Cu	-0.61	0.14	-0.18	1					
Cr	-0.14	0.32	-0.18	0.36	1				
OM	-0.61	0.29	-0.64	-0.07	0.32	1			
Clay	0.50	-0.04	0.40	0.21	0.14	-0.71	1		
Silt	0.00	-0.11	-0.14	-0.46	-0.50	0.18	-0.79*	1	
Sand	-0.46	0.07	-0.64	0.04	-0.18	0.50	-0.07	-0.32	1

Note: *significant level of $p<0.05$

Enrichment factors and index geoaccumulation were used to determine the differences of metals from anthropogenic activity and natural sources. In the present study, the trace elements considered have been lower than the EF value and I_{geo} value.

Table 9
Enrichment factors (EF) and Index of geoaccumulation (I_{geo}) of trace metals at location in KongKong Laut River

	Zn		Cd		Fe		Cu		Cr	
Site	EF	I_{geo}	EF	I_{geo}	EF	I_{geo}	EF	I_{geo}	EF	I_{geo}
S1	0.8	0.14	0.9	0.15	N/A	0.18	0.6	0.10	0.2	0.04
S2	1.0[a]	0.17	0.9	0.03	N/A	0.16	0.5	0.09	0.2	0.02
S3	1.2[a]	0.19	1.5[a]	0.23	N/A	0.15	0.6	0.09	0.3	0.04
S4	1.0[a]	0.16	0.8	0.13	N/A	0.16	0.4	0.06	0.2	0.03

Notes: [a] minor enrichment; N/A = not calculated

Table 9 shows EF and I_{geo} values from four locations along the KongKong Laut River. All four elements show the EF value under the group of the depletion to minimal enrichment. Although all elements show depletion to minimal enrichment, the EF value for Cd at S3 is higher than that of the other stations. This might be due to the enrichment of these elements coming from natural and anthropogenic sources, such as the restaurant, raft house and jetty near to the sampling location. The I_{geo} value for Fe in all four stations was shown to be close to zero, indicating that Fe behaved conservatively in the estuarine area, and therefore could be used for normalization in EF analysis (Williams & Block, 2015). As shown in Table 8, the I_{geo} value for trace metal concentration from the four sampling locations at KongKong Laut River are practically in the class of uncontaminated to moderately contaminated.

CONCLUSIONS

From the data gained, all trace metal concentrations in this study have been shown to be lower than the average shale concentration, lower than the effects-range low (ERL), and also lower than the effects-range median (ERM) parameter for aquatic toxicity. Although the data of trace metals in this study had shown low levels of concentration, continuous development without proper management might contribute to the elevation of trace metals in the study area, therefore having an impact on the aquatic ecosystem and food chain. This information could be useful, serving as a baseline for evaluating the potential impacts of future development in the area.

ACKNOWLEDGEMENTS

This research was funded by the Ministry of Higher Education and Universiti Putra Malaysia. The project was jointly supported by FRGS 2017-1 and the Putra Graduate Initiative (IPS) vot no. 9459100. The authors would like to thank the technical staff of the Faculty of Environmental Studies (FPAS) for their assistance throughout the study period.

REFERENCES

Al-Mur, B. A., Quicksall, A. N., & Al-Ansari, A. M. A. (2017). Spatial and temporal distribution of heavy metals in coastal core sediments from the Red Sea, Saudi Arabia. *Oceanologia*, *59*(3), 262-270.

Al-Najjar, T. (2011). Heavy metals pollution in sediment cores from the Gulf of Aqaba, Red Sea. *Natural Science, 3*(9), 775-782.

Al-Trabulsy, H. A. M., Khater, A. E. M., & Habbani, F. I. (2013). Heavy metals concentrations, physicochemical characteristics and natural radionuclides levels along the Saudi coastline of the Gulf of Aqaba. *Arabian Journal of Chemistry*, *6*(2), 183-189.

Cho, J., Hyun, S., Han, J. H., Kim, S., & Shin, D. H. (2015). Historical trend in heavy metal pollution in core sediments from the Masan Bay, Korea. *Marine Pollution Bulletin, 95*(1), 427-432.

Daskalakis, K. D. O., & O'Connor, T. P. (1995). Normalization and elemental sediment contamination in the coastal United States. *Environmental Science & Technology*, *29*(2), 471-477.

Gee, G. W., & Baunder, J. W. (1986). Particle size analysis. In A. Klute, A. (Ed.) *Methods of soil analysis: Part 1 (Physical and mineralogical methods)* (pp. 383-411). Madison, Wisconsin: American Society of Agronomy.

Laluraj, C. M., & Nair, S. M. (2006). Geochemical index of trace metals in the surficial sediments from the western continental shelf of India, Arabian Sea. *Environmental Geochemistry and Health*, *28*(6), 509-518.

Lim, W. Y., Aris, A. Z., & Zakaria, M. P. (2012). Spatial variability of metals in surface water and sediment in Langat River and geochemical factors that influence their water-sediment interactions. *The Scientific World Journal, 2012*, 1-14.

Long, E., Macdonald, D., Smith, S., & Calder, F. (1995). Incidence of adverse biological effects within ranges of chemical concentration in marine and estuarine sediments. *Environmental Management*, *19*(1), 81-97.

Muller, G. (1969). Index of geoaccumulation in sediments of the Rhine River. *Journal of Geology*, *2*(1969), 108-118.

Naji, A., & Ismail, A. (2011). Assessment of metals contamination in Klang River surface sediments by using different indexes. *Environment Asia*, *4*(1), 30-38.

Saraee, K. R. E., Abdi, M. R., Naghavi, K., Saion, E., Shafaei, M. A., & Soltani, N. (2011). Distribution of heavy metals in surface sediments from the South China Sea ecosystem, Malaysia. *Environmental Monitoring and Assessment*, *183*(1-4), 545-554.

Stoffers, P., Glasby, G. P., Wilson, C. J., Davis, K. R., & Walter, P. (1986). Heavy metal pollution in Wellington Harbour. *New Zealand Journal of Marine and Freshwater Research, 20*(3), 495-512.

Turekian, K. K., & Wedepohl, K. H. (19561). Distribution of the elements in some major units of the earth's crust. *Geological Society of America Bulletin*, *72*, 175-192.

Wan, E. C. K., & Mohamat-Yusuff, F. (2014). Contamination of trace elements (Cu, Pb, Cr) in Kong Ko Laut, Johor, Malaysia. In A. Z. Aris, T. H. Tengku Ismail, R. Harun, A. M. Abdullah, & M. Y. Ishak (Eds.). *From sources to solution* (pp. 567-572). Singapore: Springer.

White, S. L., & Rainbow, P. S. (1985). On the metabolic requirement for copper and zinc in molluscs and crustaceans. *Marine Environmental Research*, *16*(3), 215-229.

Williams, N., & Block, K. A. (2015). Spatial and vertical distribution of metals in sediment cores from Rio Espiritu Santo estuary, Puerto Rico, United States. *Marine Pollution Bulletin*, *100*(1), 445-452.

Zulkifli, S. Z., Ismail, A., Mohamat-Yusuff, F., Arai, T., Ismail, A, & Miyazaki, N. (2010). Johor Strait as a hotspot for trace elements contamination in Peninsular Malaysia. *Bulletin of Environmental Contamination and Toxicology*, *84*(5), 568-573.

Pertanika J. Sci. & Technol. 27 (4): 1941 - 1954 (2019)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Concentration of Heavy Metals in Street Dust and Surface Soils in Urban and Peri-Urban Regions in the Kuala Lumpur Metropolitan Area

Ahmad Ismail*, Munirah Hanapiah, Syaizwan Zahmir Zulkifli and Hishamuddin Omar and Nur Amiera Kamarudin

Department of Biology, Faculty of Science, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia

ABSTRACT

The present study aims to determine the level of heavy metals in surface soils and street dust from selected urban and peri-urban locations in the metropolitan city of Kuala Lumpur. Samples were collected from 15 different locations, where Kuala Lumpur City Centre (KLCC) served as the centre point while other locations were located at specific distances from the centre. Surface soils and street dust were collected to detect the level of contamination based on five elements (Zn, Cu, Pb, Cd and Fe). Results indicated that the metal distribution displayed a descending trend as follows: [Fe] > [Zn] > [Pb] > [Cu] > [Cd] and [Fe] > [Zn] > [Cu] > [Pb] > [Cd] in the surface soil and street dust samples, respectively. Geoaccumulation index (Igeo), contamination factor (CF) and pollution load index (PLI) were calculated. Based on the Igeo index, Fe in both samples was found to cause no pollution in all study sites. In surface soils, CF showed all sites were contaminated with Cd in both soil and street dust. In conclusion, the soil sample was polluted with metals while no pollution was found in the street dust in all sites sampled. Further intensive studies are recommended in order to determine other factors which may cause contamination.

Keywords: City centre, heavy metals, peri-urban, street dust, surface soil

ARTICLE INFO

Article history:
Received: 2 December 2018
Accepted: 30 January 2019
Published: 21 October 2019

E-mail addresses:
aismail@upm.edu.my (Ahmad Ismail)
munirahhanapiah91@gmail.com (Munirah Hanapiah)
syaizwan@upm.edu.my (Syaizwan Zahmir Zulkifli)
hishamom@upm.edu.my (Hishamuddin Omar)
nuramiera2510@gmail.com (Nur Amiera Kamarudin)
*Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

INTRODUCTION

As one of the fastest growing cities in Southeast-Asia (Asean Up, 2017; Wilson, 2018), the Kuala Lumpur metropolitan area is proudly structured with a landscape of skyscrapers, as well as steel and glass buildings. Its geographical area is occupied with industrial/ commercial activities, traffic load and development in a number of residential areas. It is no surprise that this global city has been categorized under high Human Development Index (HDI) with an estimation of 1.76 million people in 2016 within a 243 km^2 area (Anonymous, 2017; Department of Statistics Malaysia [DOSM], 2016). Consequently, undergoing projects in developing areas, or the so-called peri-urban areas, are vital in ensuring that the human capacity is parallel with the social and economic development in Kuala Lumpur.

While the acceleration of development occurs, the issue that comes to light is the growing repercussions of the human activities in developed areas (Janaydeh et al., 2016). The pollutant discharge from the urbanization process to the environmental surroundings is believed to cause a significant threat to mankind (Christoforidis & Stamatis, 2009; Udechukwu et al., 2015). One of the chief pollutants in urban areas nowadays is heavy metal.

To deal with this pollutant, it is important to know the sources of the pollutant, and the type of metals found in that particular area. Furthermore, a collection of street dust and surface soils are believed to give the reflection of environmental conditions (Yu et al., 2012) by assessing the health hazard associated with the availability of particular metals. Moreover, previous studies have revealed that metals from the samples-are a reflection of street dust and soil samples around the urban area which were expected from urban emissions (e.g. traffic emissions, industrial discharges, domestic heating, waste incineration, and many more) (Chen et al., 2005; Christoforidis & Stamatis, 2009; Li et al., 2001), while, the activities resulting from agricultural land is the main source of metals in rural areas (Chen et al., 2016). However, there is limited information regarding the level and sources of pollution in peri-urban areas.

Hence, the present study aims to determine the concentration of selected metals (Zn, Fe, Cd, Pb and Cu) in street dust and surface soil samples and to evaluate the interconnection of metals released in urban areas, which is affecting the availability of metals in peri-urban metropolitan areas of Kuala Lumpur.

MATERIALS AND METHODS

Study Area Descriptions and Samples Collection

Located in Peninsular Malaysia, the Federal Territory of Kuala Lumpur is one of the 14 states of Malaysia (Figure 1). In this metropolitan city, the KLCC skyscraper has been chosen as the centre for location sampling as this structure proudly represents the urbanisation of Kuala Lumpur. Hence, about 15 specific locations were chosen within the KL metropolitan area which was based on a 2, 4, and 8 km radius ring from the city centre, KLCC (Figure 2) including the peri-urban areas within this city. All the 15 locations of the study areas are described in Table 1.

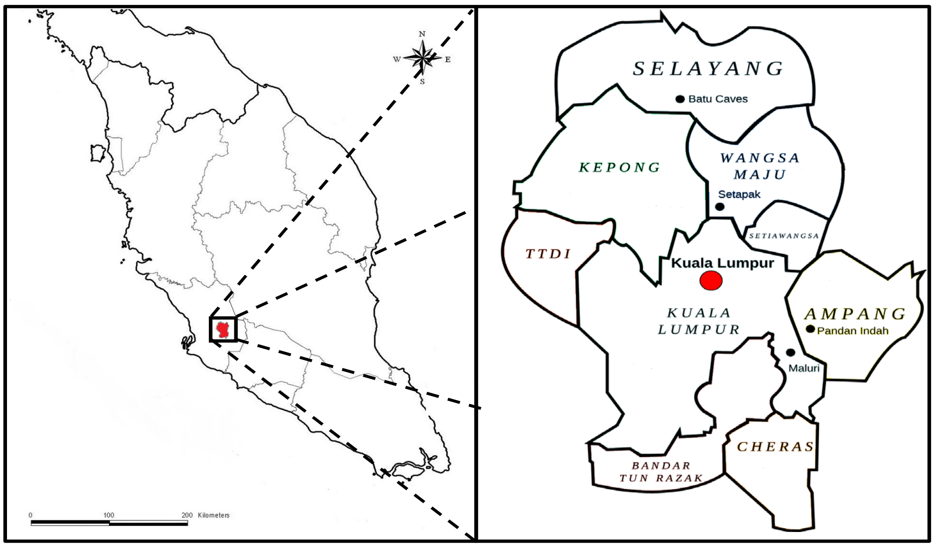


Figure 1. Kuala Lumpur metropolitan area

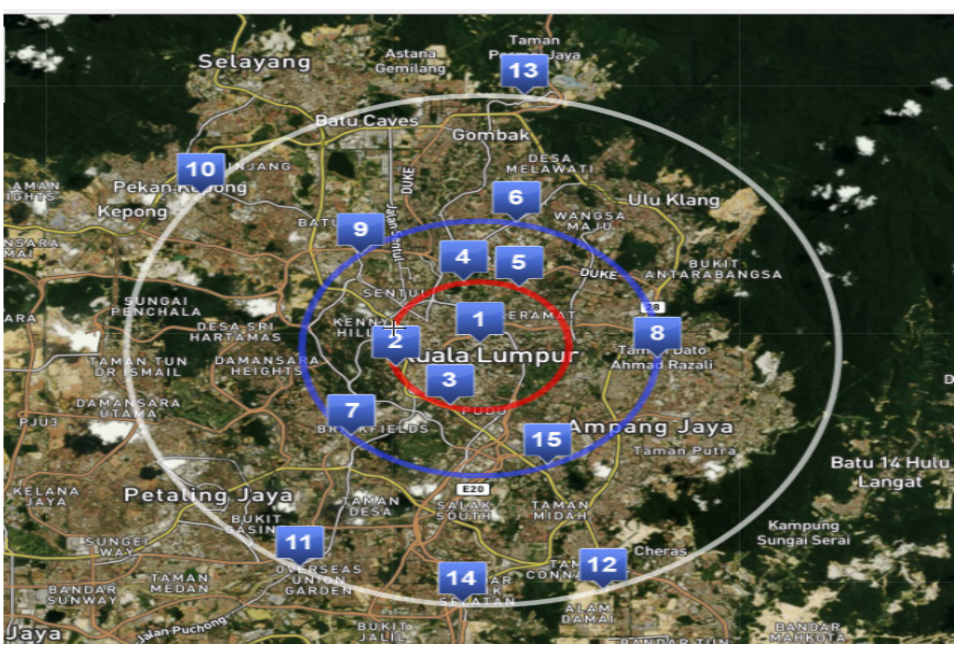


Figure 2. Radius ring of the 15 locations within the study area in Kuala Lumpur

Table 1

Description of each study site

Site no.	Geographic locations (N,E)	Description area	Traffic load	Land uses
1	3.159162, 101.711182	Kuala Lumpur City Centre (KLCC)	Heavy	Commercial area, Tourism area
2	3.150712, 101.689948	Parlimen Road	Moderate	Commercial area
3	3.137101, 101.703295	Imbi Road	Heavy	Commercial area
4	3.176986, 101.710896	Temerloh Road	Low	Recreational area
5	3.179541, 101.721193	Sultan Yahya Petra Road	Heavy	Commercial area, Residential area, Industrial area
6	3.201768, 101.720317	Genting-Klang Road	Heavy	Commercial area, Residential area
7	3.126773, 101.678706	Bangsar Road	Heavy	Commercial area, Residential area
8	3.153396, 101.757273	Ampang Road	Moderate	Residential area
9	3.19038, 101.680763	Sultan Azlan Shah Road	Moderate	Residential area, Commercial area
10	3.214067, 101.641241	Kepong Road	Heavy	Commercial area, Residential area
11	3.079295 101.665211	Puchong Road	Heavy	Commercial area, Residential area
12	3.068612, 101.743021	Alam Damai Road	Moderate	Commercial area
13	3.243951, 101.720619	Gombak Road	Moderate	Commercial area, Residential area
14	3.066855, 101.706701	Sungai Besi Expressway (SBE)	Heavy	Commercial area
15	3.115239, 101.727888	Cheras Road	Heavy	Commercial area, Residential area

Street dusts were collected through leaf collection from roadside trees (Moreno et al., 2003; Ram et al., 2014; Tanushree et al., 2011) and a scoop was used in order to collect the surface soil at every study site. Surface soils were collected within a 5cm depth from the top of the roadside soil area. Both samples were brought back to the laboratory for subsequent preparation and analysis.

Samples Preparation and Analysis

Collected leaves were divided into two categories, washed and unwashed. Then, both types of samples (leaf and soil) were dried in an oven for approximately 2-3 days at 60˚C temperature, for the purpose of removing trapped moisture. After that, the leaves sample were cut/ ground into small pieces using a mortar and pestle while surface soil samples were ground and sieved prior to acid-digestion experiment. Both samples were digested based on the method described by Abubakar et al. (2018). Filtered digested samples were then analysed using the Atomic Absorption Spectroscopic (AAS) (Shimadzu 6800) for selected elements (Cu, Zn, Cd, Pb and Fe).

Quality Assurance and Quality Control

All the glassware and plastic ware were soaked with 10 ml of nitric acid for decontamination purposes. The reagent blank was prepared for each acid-digestion set. Meanwhile, samples and standard reference material (SRM 1646a) were digested in tri-replicates each. The recovery percentage for each of the selected heavy metals is satisfactory: Cu (103.63%), Zn (97.29%), Cd (90.77%), Pb (90.38%) and Fe (95.46%), respectively.

Data Analysis

Data were analysed using a statistical software, SPSS ver. 22. One-way ANOVA was performed while correlation analysis was conducted to find the possible relationship amongst the variables.

Pollution Indices

Geoaccumulation index (Igeo), Contamination factor (CF) and Pollution Load Index (PLI) were determined. All the formulas and calculations are referring to the previous studies as stated below:

i. Igeo calculation was adapted from the study proposed by Muller (1969). Hence, the background value of the metals (Cu, 48; Zn, 95; Pb, 20; Ni, 68 and Cd, 0.30) was chosen referring to Muller (1969).

ii. The Contamination Factor (CF) is an index which is expressed as the ratio of metal concentration over background value. The calculation of this index was based on the study by Abubakar et al. (2018) and Turekian and Wedepohl (1961).

iii. Meanwhile, the Pollution Load Index (PLI) is an ecological tool to assess the extent of metal pollution or numbers of metal at a particular study area (Abubakar et al., 2018; Cabrera et al., 1999).

RESULTS AND DISCUSSIONS

Concentration of Metals Content in Street Dust

The concentration of analyzed elements was determined from all 15 study sites in Kuala Lumpur metropolitan areas. Assessment of metals in street dust samples showed that metal pollution resulted from human modernization exploits. In general, element content abundance can be ranked as follows: [Fe] > [Zn] > [Cu] > [Pb] > [Cd].

Iron (Fe) is an element found highest in street dust samples. This metal's abundance possibly comes from brake dust, vehicle rust and motorcar exhaust (Garg et al., 2000; Hopke et al., 1980; Weber et al., 2000). In this study, the concentration of Fe was found to be highest at site no. 3 while lowest at site no. 2 (Table 2). As described in Table 1, site no. 3 is located at Imbi Road which is loaded with heavy traffic and commercial development areas. Besides, samples were collected at the road junction and near to LRT and KL monorail lanes. Having a huge number of vehicles on the road and located in a commercial area, are the reasons Fe element was found highest in the street dust samples. Emission of vehicles (brake dust, exhaust) happened to increase the availability of this element and thus contaminating this study area. In addition to that, site no. 2 (Parlimen Road) reflected the lowest concentration of Fe. The samples were collected near to the pavement at the roadside which is a small road, moderately loaded with vehicles. This sampled area is surrounded by green trees and grass planted along the road. Moderate in traffic load, the green surrounding helps to regulate the cycle of air and the concentration of Fe is the lowest among all study sites.

Table 2

Concentration of metals (mean ± S.D) (µg.g-1) found in the street dust samples from all 15 study sites

Sites	Zn	Cu	Pb	Cd	Fe
1	24.68 ± 1.23	6.85 ± 0.34	1.56 ± 0.08	0.20 ± 0.01	144 ± 7.20
2	4.59 ± 0.23	11.34 ± 0.57	5.76 ± 0.29	0.87 ± 0.04	2 ± 0.10
3	4.31 ± 0.22	4.30 ± 0.22	3.84 ± 0.19	0.66 ± 0.03	921 ± 46.05
4	14.37 ± 0.72	7.54 ± 0.38	1.8 ± 0.09	1.67 ± 0.08	84 ± 4.20
5	43.76 ± 2.19	7.87 ± 0.39	1.02 ± 0.05	0.26 ± 0.01	12 ± 0.60
6	113.48 ± 5.67	2.72 ± 0.14	8.82 ± 0.44	1.34 ± 0.07	123 ± 6.15
7	80.57 ± 4.03	6.39 ± 0.32	2.28 ± 0.11	1.00 ± 0.05	389 ± 19.45
8	1.43 ± 0.07	2.14 ± 0.11	0.9 ± 0.05	1.27 ± 0.06	276 ± 13.80
9	17.68 ± 0.88	4.30 ± 0.21	7.74 ± 0.39	1.15 ± 0.06	200 ± 10.00
10	1.44 ± 0.07	3.44 ± 0.17	2.58 ± 0.13	1.53 ± 0.08	262 ± 13.10
11	96.64 ± 4.83	11.60 ± 0.58	0.18 ± 0.01	1.81 ± 0.09	59 ± 2.95
12	16.49 ± 0.82	9.85 ± 0.49	1.2 ± 0.06	0.21 ± 0.01	199 ± 9.95
13	28.28 ± 1.41	1.07 ± 0.05	0.42 ± 0.02	0.59 ± 0.03	156 ± 7.80
14	21.28 ± 1.06	18.10 ± 0.91	3.66 ± 0.18	0.07 ± 0.00	260 ± 13.00
15	22.77 ± 1.14	9.08 ± 0.45	5.34 ± 0.27	1.13 ± 0.06	75 ± 3.75

Zinc (Zn) is one of the most significant pollutants found in street dust studies. Availability of this pollutant is believed to be caused by tire treads. In the vulcanizing process, zinc oxide (ZnO) acts as an activator in vehicle tire production (Adachi & Tainosho, 2004). Hence, in this study, Zn was found to be the second highest element with a high concentration at site no. 6 and lowest at site no. 8. Site no. 6 represents Genting-Klang Road which was described as a heavy traffic load area (Table 1). Increasing the number of traffic on the road and human activities in the area have increased the availability of Zn found in street dust which contributes as one of the sources of heavy metal pollution (Zheng et al., 2010).

The metal content of Cu in street dust was well assessed. This element possibly resulted from the corrosion of metallic parts of vehicles (e.g engine wear, thrust bearing, brushing and bearing metals) (Al-Khashman, 2004, 2007; Al-Khashman & Shawabkeh, 2006). The level of copper in street dust samples was high at study site no. 14. This site represents the Sungai Besi Expressway (SBE) Road which is one of the highways in Klang Valley. Traffic load is high in this area and believed to have cause the high metal loading, especially copper. Copper was found to be the lowest among all 15 sites especially for site no. 13 (Gombak Road). This area was low in copper content which might be due to the traffic load and the effects of human activities nearby.

Generally, lead (Pb) and cadmium (Cd) are low in concentration for all sites. Pb is said to be the element most concerned in heavy metal pollution (Christoforidis & Stamatis, 2009). Detection of this metal is directly associated with the emission of the exhaust vehicle and industrial emissions released. Pb metal is believed to be released from the combustion of gasoline that contains tetraethyl lead (anti-knock agent) (Al-Khashman, 2007; Janaydeh et al., 2018; Tüzen, 2003). Meanwhile, the availability of Cd metals polluted in the environment comes from combustion products (motor vehicles, carburettors) (Charlesworth et al., 2003; Shinggu et al., 2007). From the data collection, both elements have reflected the low mean concentration of elements polluting the KL metropolitan area. Based on Table 2, site no. 11 (Puchong Road) shows the highest Cd metal while low in Pb content. Site no. 6 was highest in Pb content while site no. 14 showed low levels of Cd.

Concentration of Metals Content in Surface Soil

The surface of the soil in an urban and peri-urban area has been significantly disturbed by the metal pollution through some anthropogenic emissions. All the atmospheric pollutant definitely ends up on the soil surface eventually swept by the water causing water pollution (e.g groundwater, freshwater, ocean). Mean concentration of 5 selected elements were well determined. Besides that, CRM recoveries for metals in street dust samples varied from 87-102%. The decreasing trend of assessment elements is as follows, [Fe] > [Zn] > [Pb] > [Cu] > [Cd].

Fe content showed the highest level in the soil samples at all study sites as this element was naturally abundant and can be found in high value in uncontaminated areas. However, this study found that this high concentration of Fe could be caused by human activities such as emissions and combustions. It was highest at site no. 1 (KLCC), but the lowest was at site no. 10 (Kepong Road). It is no surprise that Fe was high in topsoil at the KLCC site as seen and been predicted during its early constructions; while Kepong Road holds the lowest record for Fe.

Zn element was the highest at site no. 5 (Sultan Yahya Petra Road), while being the lowest at site no. 14 (SBE). It is due to locality of site no. 5 which is surrounded by residential areas, commercial areas and also industrial activities which is compounded with the active load of traffic on the road. This varies from the SBE site that has predicted sources of outcome from traffic emissions.

Lead metal loading in soil samples was shown highest in samples from site no. 8 (Ampang Road), lowest in the sample from site no. 13 (Gombak Road). Vehicle emission from the combustion of leaded gasoline has given the impact to the level of lead in the soil surface of the study site. Even though the traffic load is moderate in this area, this traffic lane becomes a busy road during peak hours compared to site no. 13 and this contributes to the accumulation of combustion products released to the atmosphere, which contains lead.

Table 3

Concentration of metals (mean ± S.D) (µg g-1) found in the surface soil samples from all 15 study sites

Sites	Zn	Cu	Pb	Cd	Fe*
1	257.00 ± 12.85	26.75 ± 1.34	64.14 ± 3.21	3.94 ± 0.20	15.89± 0.79
2	281.53 ± 14.07	23.35 ± 1.17	104.22 ± 5.21	4.01 ± 0.20	9.07± 0.45
3	303.48 ± 15.17	18.70 ± 0.94	74.04 ± 3.70	2.87 ± 0.14	12.72± 0.64
4	264.76 ± 13.23	29.75 ± 1.49	166.62 ± 8.33	3.94 ± 0.20	8.21± 0.41
5	81.17 ± 4.05	8.14 ± 0.41	41.34 ± 2.07	3.80 ± 0.19	8.17± 0.41
6	298.94 ± 14.94	31.00 ± 1.55	49.74 ± 2.49	3.99 ± 0.20	12.69± 0.63
7	254.32 ± 12.71	36.45 ± 1.82	123.84 ± 6.19	2.80 ± 0.14	11.92± 0.60
8	366.80 ± 18.34	67.10 ± 3.36	312.00 ± 15.60	4.47 ± 0.22	10.25± 0.51
9	299.48 ± 14.97	94.65 ± 4.73	130.50 ± 6.53	3.74 ± 0.19	6.78± 0.34
10	200.76 ± 10.03	17.37 ± 0.87	68.76 ± 3.44	3.87 ± 0.19	3.88± 0.19
11	143.69 ± 7.18	58.86 ± 2.94	195.48 ± 9.77	3.00 ± 0.15	13.42± 0.67
12	262.38 ± 13.11	11.03 ± 0.55	110.34 ± 5.52	4.41 ± 0.22	10.22± 0.51
13	237.35 ± 11.86	19.02± 0.95	37.02 ± 1.85	3.52 ± 0.18	8.07± 0.40
14	452.88 ± 22.64	91.60 ± 4.58	111.24 ± 5.56	3.40 ± 0.17	4.95± 0.25
15	282.96 ± 14.14	51.55 ± 2.58	113.94 ± 5.70	3.87 ± 0.19	5.25± 0.26

Note: Concentration of Fe is in mg.g-1

In soil samples, out of 15 study sites, copper showed the highest concentration at site no. 9, while lowest at site no. 5 (Table 3). Site no. 9 represents Sultan Azlan Shah Road. It is described as an area with moderate traffic load, surrounded by residential and commercial areas. Ongoing construction in this area has played a vital role in contributing to the input of copper to the soil surroundings even though the main sources of copper are believed to come from vehicle parts. Cd had the least concentration of an element found from all 15 sites. Data showed that Cd element was similar for all study sites (Table 3).

Correlation Analysis amongst Heavy Metal Elements in Street Dusts and Surface Soils

The relationship between metals in street dust and soil surface samples were investigated (Table 4). In street dust samples, mostly, all the metal elements have shown a negative relationship and a small positive relationship value amongst each other. Meanwhile, in surface soil samples, there is a moderate positive correlation between elements Cu and Zn (0.652) and Pb and Cu (0.494) while the other elements showed a slightly positive relationship. This shows that the distribution of Cu pollutant in soil samples is related to the presence of Pb and Zn, as the main sources of these three elements are mainly sources from the vehicle usages on the road.

Table 4

Interelement correlations for street dust and surface soil samples from the whole study areas

Elements	Zn	Cu	Pb	Cd	Fe
Street dust					
Zn	1.000				
Cu	-0.137	1.000			
Pb	-0.170	-0.194	1.000		
Cd	0.242	-0.005	0.041	1.000	
Fe	-0.008	-0.233	0.129	0.052	1.000
Surface Soil					
Zn	1.000				
Cu	0.652	1.000			
Pb	0.199	0.494	1.000		
Cd	-0.028	-0.133	0.168	1.000	
Fe	-0.158	-0.224	0.061	-0.132	1.000

Note: Significant level, p-value (p=0.05)

Assessment of Metals Contamination in Samples

Assessment of element contamination in street dust samples was evaluated (Table 5). From the Igeo assessment, all the 15 study sites with the presence of these elements (Zn, Cu, Pb and Fe) were recognized as unpolluted areas. Meanwhile, Cd with the lowest value of element concentration has recorded moderate to strongly polluted areas; whereas site no. 11 was reported to be strongly polluted, site nos. 4, 6, 7, 8, 9, 10 moderately polluted and while areas categorized as unpolluted to moderately polluted where found at study sites 2, 3 and 13. By referring to the contamination factor (CF), street dusts in KL metropolitan area had low contamination with respect to Cu, Pb, Fe and Zn (except at site no. 6 and 11); moderate contamination with Cd at site 2, 3 and 13; considerable contamination at site 4, 6, 7, 8, 9, 10 and 15; while very high contamination of Cd at site no. 11. Another assessment was evaluated using the pollution load index (PLI). Overall, all the areas of Kuala Lumpur were categorized as areas with no metal pollution.

However, in surface soil samples, it does reflect a different evaluation (Table 6). For Igeo calculation, all 15 sites were unpolluted with Fe, Cu (except site no. 9, 14) and Zn at the site no. 5. While Cd element was strongly polluted at study site no. 1, 2, 4, 5, 6, 8, 9, 10, 12 and 15. For the next assessment, CF generally showed that there were no contamination of Fe at all study sites, and no contamination of Cu at most study sites. All sites were very highly contaminated with Cd, Pb for selected sites (4, 7, 8, 9, 11). PLI concluded that all study sites were categorized as polluted areas with the presence of heavy metals in surface soils.

Table 5

Summary of Igeo, CF and PLI for street dust samples in 15 sampling sites

	I_{geo}					CF					PLI
Sites	Zn	Cu	Pb	Cd	Fe	Zn	Cu	Pb	Cd	Fe	
1	-2.53	-3.301	-4.265	-1.184	-8.94	0.2598	0.152	0.078	0.660	0.0031	0.091
2	-4.96	-2.574	-2.381	0.945	-15.11	0.0483	0.252	0.288	2.888	0.0000	0.053
3	-5.05	-3.972	-2.966	0.553	-6.26	0.0454	0.096	0.192	2.200	0.0195	0.129
4	-3.31	-3.162	-4.059	1.889	-9.72	0.1513	0.168	0.09	5.555	0.0018	0.118
5	-1.70	-3.101	-4.878	-0.769	-12.53	0.4606	0.175	0.051	0.880	0.0003	0.062
6	-0.33	-4.636	-1.766	1.570	-9.17	1.1945	0.060	0.441	4.455	0.0026	0.206
7	-0.82	-3.401	-3.718	1.149	-7.51	0.8482	0.142	0.114	3.328	0.0082	0.207
8	-6.64	-4.979	-5.059	1.497	-8.00	0.0151	0.048	0.045	4.235	0.0058	0.060
9	-3.01	-3.974	-1.955	1.350	-8.47	0.1862	0.095	0.387	3.823	0.0042	0.162
10	-6.63	-4.297	-3.540	1.770	-8.08	0.0152	0.076	0.129	5.115	0.0056	0.084
11	-0.56	-2.541	-7.381	2.005	-10.23	1.0173	0.258	0.009	6.023	0.0013	0.112
12	-3.11	-2.777	-4.644	-1.127	-8.47	0.1736	0.219	0.06	0.687	0.0042	0.092
13	-2.33	-5.979	-6.158	0.401	-8.83	0.2977	0.024	0.021	1.980	0.0033	0.063
14	-2.74	-1.899	-3.035	-2.599	-8.09	0.2241	0.402	0.183	0.248	0.0055	0.118
15	-2.65	-2.894	-2.490	1.328	-9.88	0.2397	0.202	0.267	3.767	0.0016	0.151

Table 6

Summary of Igeo, CF and PLI for surface soil samples in 15 sampling site

	I_{geo}					CF					PLI
Sites	Zn	Cu	Pb	Cd	Fe	Zn	Cu	Pb	Cd	Fe	
1	0.85	-1.34	1.096	3.128	-2.156	2.705	0.594	3.207	13.117	0.337	4.592
2	0.98	-1.53	1.797	3.155	-2.965	2.964	0.519	5.211	13.363	0.192	4.596
3	1.09	-1.85	1.303	2.674	-2.477	3.195	0.416	3.702	9.570	0.269	4.472
4	0.89	-1.18	2.474	3.128	-3.108	2.787	0.661	8.331	13.117	0.174	4.729
5	-0.81	-3.05	0.463	3.079	-3.115	0.854	0.181	2.067	12.677	0.173	2.674
6	1.07	-1.12	0.729	3.149	-2.480	3.147	0.689	2.487	13.310	0.269	5.340
7	0.84	-0.89	2.045	2.635	-2.570	2.677	0.810	6.192	9.320	0.253	4.721
8	1.36	-0.01	3.379	3.313	-2.788	3.861	1.491	15.600	14.903	0.217	5.890
9	1.07	0.49	2.121	3.054	-3.384	3.152	2.103	6.525	12.457	0.144	6.595
10	0.49	-1.96	1.197	3.104	-4.190	2.113	0.386	3.438	12.897	0.082	4.152
11	0.01	-0.20	2.704	2.738	-2.399	1.513	1.308	9.774	10.010	0.284	3.494
12	0.88	-2.61	1.879	3.291	-2.792	2.762	0.245	5.517	14.685	0.217	3.896
13	0.74	-1.83	0.303	2.968	-3.133	2.498	0.423	1.851	11.733	0.171	3.690
14	1.67	0.44	1.891	2.917	-3.838	4.767	2.036	5.562	11.330	0.105	6.764
15	0.99	-0.39	1.925	3.104	-3.754	2.979	1.146	5.697	12.897	0.111	5.645

CONCLUSIONS

This study revealed that soil samples from all study sites were polluted with metals while, street dust did not indicate any signs of metal pollution in the surrounding atmosphere. Thus, at this moment, the distribution of metals through emissions from human activities in the Kuala Lumpur metropolitan area is not influenced by the urban areas around KLCC; based on the metal concentration recorded in street dust samples from all locations. However, in all the locations within the KL metropolitan area showing moderate to heavy traffic load and moderate to large human capacity, the presence of metal pollutants in street dust and soil samples varied depending on the surroundings. As a suggestion, the continuous assessment of metal concentration in street dust samples in the KL metropolitan area is encouraged to ensure that the data recorded is up-to-date. Meanwhile, an immediate response would be to carry out surface soil cleaning as to remediate the metal contaminated soils.

ACKNOWLEDGEMENT

This study was supported by Putra Grant - Putra Group Initiative (IPB) (Reference no.: GP-IPB/2014/9433801) funded by Universiti Putra Malaysia.

REFERENCES

Abubakar, U. S., Zulkifli, S. Z., & Ismail, A. (2018). Heavy metals bioavailability and pollution indices evaluation in the mangrove surface sediment of Sungai Puloh, Malaysia. *Environmental Earth Sciences, 77*(6), 225-237.

Adachi, K., & Tainosho, Y. (2004). Characterization of heavy metal particles embedded in tire dust. *Environment International, 30*(8), 1009-1017.

Al-Khashman, O. A. (2004). Heavy metal distribution in dust, street dust and soils from the work place in Karak Industrial Estate, Jordan. *Atmospheric Environment, 38*(39), 6803-6812.

Al-Khashman, O. A. (2007). The investigation of metal concentrations in street dust samples in Aqaba city, Jordan. *Environmental Geochemistry and Health, 29*(3), 197-207.

Al-Khashman, O. A., & Shawabkeh, R. A. (2006). Metals distribution in soils around the cement factory in southern Jordan. *Environmental Pollution, 140*(3), 387-394.

Anonymous. (2017b). *Kuala Lumpur population.* Retrieved July 22, 2016, from http://worldpopulationreview.com/world-cities/kuala-lumpur

Asean Up. (2017). *Infographic: Top cities and urbanization in ASEAN.* Retrieved April 3, 2018, from https://aseanup.com/infographic-top-cities-urbanization-asean/.

Cabrera, F., Clemente, L., Barrientos, E. D., López, R., & Murillo, J. M. (1999). Heavy metal pollution of soils affected by the Guadiamar toxic flood. *Science of the Total Environment, 242*(1-3), 117-129.

Charlesworth, S., Everett, M., McCarthy, R., Ordonez, A., & De Miguel, E. (2003). A comparative study of heavy metal concentration and distribution in deposited street dusts in a large and a small urban area: Birmingham and Coventry, West Midlands, UK. *Environment International, 29*(5), 563-573.

Chen, T. B., Zheng, Y. M., Lei, M., Huang, Z. C., Wu, H. T., Chen, H., … Tian, Q. Z. (2005). Assessment of heavy metal pollution in surface soils of urban parks in Beijing, China. *Chemosphere, 60*(4), 542-551.

Chen, T., Chang, Q., Liu, J., Clevers, J. G. P. W., & Kooistra, L. (2016). Identification of soil heavy metal sources and improvement in spatial mapping based on soil spectral information: A case study in northwest China. *Science of the Total Environment, 565*(2016), 155-164.

Christoforidis, A., & Stamatis, N. (2009). Heavy metal contamination in street dust and roadside soil along the major national road in Kavala's region, Greece. *Geoderma, 151*(3-4), 257-263.

Department of Statistics Malaysia. (2016). *Federal Territory of Kuala Lumpur.* Retrieved April 05, 2018 from https://www.dosm.gov.my/v1/index.php?r=column/cone&menu_id=bjRlZXVGdnBueDJKY1BPWEFPRlhIdz09

Garg, B. D., Cadle, S. H., Mulawa, P. A., Groblicki, P. J., Laroo, C., & Parr, G. A. (2000). Brake wear particulate matter emissions. *Environmental Science & Technology, 34*(21), 4463-4469.

Hopke, P. K., Lamb, R. E., & Natusch, D. F. (1980). Multielemental characterization of urban roadway dust. *Environmental Science & Technology, 14*(2), 164-172.

Janaydeh, M., Ismail, A., Omar, H., Zulkifli, S. Z., Bejo, M. H., & Aziz, N. A. A. (2018). Relationship between Pb and Cd accumulations in house crow, their habitat, and food content from Klang area, Peninsular Malaysia. *Environmental Monitoring and Assessment, 190*(1), 47-62.

Janaydeh, M., Ismail, A., Zulkifli, S. Z., Bejo, M. H., Aziz, N. A. A., & Taneenah, A. (2016). The use of feather as an indicator for heavy metal contamination in house crow (Corvus splendens) in the Klang area, Selangor, Malaysia. *Environmental Science and Pollution Research, 23*(21), 22059-22071.

Li, X., Poon, C. S., & Liu, P. S. (2001). Heavy metal contamination of urban soils and street dusts in Hong Kong. *Applied Geochemistry, 16*(11-12), 1361-1368.

Moreno, E., Sagnotti, L., Dinarès-Turell, J., Winkler, A., & Cascella, A. (2003). Biomonitoring of traffic air pollution in Rome using magnetic properties of tree leaves. *Atmospheric Environment, 37*(21), 2967-2977.

Muller, G. (1969). Index of geoaccumulation in sediments of the Rhine River. *Geojournal, 2*(1969), 108-118.

Ram, S. S., Kumar, R. V., Chaudhuri, P., Chanda, S., Santra, S. C., Sudarshan, M., & Chakraborty, A. (2014). Physico-chemical characterization of street dust and re-suspended dust on plant canopies: An approach for finger printing the urban environment. *Ecological Indicators, 36*(2014), 334-338.

Shinggu, D. Y., Ogugbuaja, V. O., Barminas, J. T., & Toma, I. (2007). Analysis of street dust for heavy metal pollutants in Mubi, Adamawa State, Nigeria. *International Journal of Physical Sciences, 2*(11), 290-293.

Tanushree, B., Chakraborty, S., Bhumika, F., & Piyal, B. (2011). Heavy metal concentrations in street and leaf deposited dust in Anand city, India. *Research Journal of Chemical Sciences, 1*(5), 61-66.

Turekian, K. K., & Wedepohl, K. H. (1961) Distribution of the elements in some major units of the Earth's crust. *The Geological Society of America Bulletin, 72*(2), 175-192.

Tüzen, M. (2003). Determination of heavy metals in soil, mushroom and plant samples by atomic absorption spectrometry. *Microchemical Journal, 74*(3), 289-297.

Udechukwu, B. E., Ismail, A., Zulkifli, S. Z., & Omar, H. (2015). Distribution, mobility, and pollution assessment of Cd, Cu, Ni, Pb, Zn, and Fe in intertidal surface sediments of Sg. Puloh mangrove estuary, Malaysia. *Environmental Science and Pollution Research, 22*(6), 4242-4255.

Weber, S., Hoffmann, P., Ensling, J., Dedik, A. N., Weinbruch, S., Miehe, G., Gutlich, P., & Ortner, H. M. (2000). Characterization of iron compounds from urban and rural aerosol sources. *Journal of Aerosol Science, 31*(8), 987-997.

Wilson, R. (2018). *The rush for rail in Southeast Asia's fastest growing cities.* Retrieved July 28, 2018, from https://www.jllrealviews.com/economy/infrastructure/rush-rail-southeast-asias-fastest-growing-cities/

Yu, S., Zhu, Y. G., & Li, X. D. (2012). Trace metal contamination in urban soils of China. *Science of the Total Environment, 421*(2012), 17-30.

Zheng, N., Liu, J., Wang, Q., & Liang, Z. (2010). Health risk assessment of heavy metal exposure to street dust in the zinc smelting district, Northeast of China. *Science of the Total Environment, 408*(4), 726-733.

Pertanika J. Sci. & Technol. 27 (4): 1955 - 1959 (2019)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Long-Term Water Quality Assessment in a Tropical Monsoon

Jabir Harna Abdulkareem[1,2], Da'u Abba Umar[1], Alhassan Idris Gabasawa[2], Chinedum Anyika[3] and Nor Rohaizah Jamil[1]*

[1]Department of Environmental Sciences, Faculty of Environmental Studies, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia
[2]Department of Soil Science, Institute for Agricultural Research/Faculty of Agriculture, Ahmadu Bello University, P.M.B 1044, Zaria, Nigeria
[3]Faculty of Biosciences and Medical Engineering, Universiti Teknologi Malaysia, 81310 Skudai, Johor Bahru, Malaysia

ABSTRACT

Multivariate statistical techniques such as principal component analysis (PCA) and cluster analysis (CA) were applied to water quality parameters in order to interpret complex matrices for better assessment of water quality and environmental status of a watershed. A study was conducted to assess water quality and to establish relationship among water quality parameters in Kelantan River basin. Water quality data was obtained from Department of Environment, (DOE) Malaysia from 2005-2014. Multivariate statistical techniques such as principal component analysis (PCA) and cluster analysis (CA) were applied to 15 water quality parameters in order to interpret complex matrices for better assessment of water quality and environmental status of the watershed. From the results, five PCs were extracted which are collectively accountable for controlling approximately 70% of the watershed's water quality. Results of cluster analysis indicated that three water quality parameters that included total suspended solids, total solids and turbidity control the water quality of the study area. These parameters were allocated into three clusters based on their similarity. The finding of this study will contribute to existing knowledge of the problems

ARTICLE INFO

Article history:
Received: 3 December 2018
Accepted: 30 January 2019
Published: 21 October 2019

E-mail addresses:
abdulkareemjabir@yahoo.com (Jabir Haruna Abdulkareem)
daumarkukuma@gmail.com (Da'u Abba Umar)
algabasiwwu@yahoo.com (Alhassan Idris Gabasawa)
anyika3c@yahoo.com (Chinedum Anyika)
norrohaizah@upm.edu.my (Nor Rohaizah Jamil)
* Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

associated with water quality in the basin. This information can be put to use by land use managers and policy makers for future planning and development of the watershed.

Keywords: Cluster analysis, Kelantan, multivariate statistics, principal component analysis, water quality

INTRODUCTION

Water quality in Kelantan River basin has been under threat in the past few decades due to land use changes such as unrestricted deforestation for logging activities, agricultural activities and urbanization (Abdulkareem et al., 2017). In addition to this, natural factors such as soil type, geology, erosion, weathering of crustal materials, topography, vegetation cover, rainfall (intensity, duration and amount) have been reported to cause seasonal variation in water quality of the basin. While natural factors such as soil type, topography and rainfall are stable over a certain period of time, anthropogenic factors such as land use vary from time to time due to rapid increase in urbanization as well as for sustainable watershed management. The spatio-temporal differences existing in physicochemical properties make it necessary to have representative measurements of water quality parameters. However, monitoring activities are designed to cover periodic water samples collection at several locations aimed at determining several parameters which usually lead to collection of large data set that require complex interpretation (Simeonov et al., 2003; Yu et al., 2013).

Multivariate statistical techniques such as principal component analysis (PCA) and cluster analysis (CA) can be applied to water quality parameters in order to interpret complex matrices for better assessment of the chemistry and environmental status of the watershed. They can also be used in pinpointing sources controlling water quality and provide possible water management techniques as well as offers speedy response to pollution problems (Jalali, 2010). Several researchers in the past have utilized the use of multivariate statistical analysis in analyzing water quality data. For example, Dalakoti et al. (2018) used multivariate analyses (HCA [hierarchical cluster analysis] and PCA) to examine the potential similarities in pollution loads and the factors responsible for pollution at Nainital District in India. In Malaysia, Saiful et al. (2017) identified various types of pollution sources due to changes in land uses in Perlis River basin using HCA, PCA and CA. Multivariate analyses was applied in Lake Victoria, Kenya to evaluate the use of changes in pollution indicators (Kundu et al., 2017). The spatial variation and potential sources of pollution were identified by Juahir et al. (2011) using hierarchical agglomerative cluster analysis (HACA), PCA and CA in Langat River basin, Malaysia. Dalakoti et al. (2018) analyzed water quality data at wetland monitoring stations in South Florida, USA. They applied multivariate analysis with the aim of identifying variance in water depth and water quality variables due to changes in rainfall seasons for both wet and dry season. In another study, PCA and CA were utilized to assess spatial variations and to interpret water quality data of Lis River water, Portugal (Vieira et al., 2012). Different multivariate statistical techniques (CA, PCA and multiple regression) were applied by Simeonov et al. (2003) to datasets obtained from northern Greece for the assessment of surface water quality during a monitoring program.

In this study, Kelantan River basin was selected due to its flood prone nature in Malaysia. Several factors are responsible for flood in the watershed such as mismanaged drainage system, unpredictable nature of weather conditions and unplanned development by human activities. LULC changes and climate change have significant impact on the natural hydrologic conditions and ecological process of the watershed. LULC changes increase the occurrence of flooding, presenting a significant management problem. Furthermore, extensive LULC changes recorded 1980s to 2000s, especially in relation to deforestation (due to logging activities) and transformation to agricultural land (mostly for rubber and oil palm production) have been reported by several authors (Adnan & Atkinson, 2011; Abdulkareem et al., 2018a; Abdulkareem et al., 2018b; Jamaliah, 2007; Wan, 1996). In view of this, this work was carried out to assess water quality and to establish a relationship among water quality parameters.

MATERIALS AND METHODS

Monitoring Area

Kelantan River basin is in the northeastern part of Peninsular Malaysia between latitudes 4° 40' and 6° 12' north, and longitudes 101° 20' and 102° 20' East. The catchment has a maximum length and breadth of about 150 km and 140 km respectively. The watershed has an area of approximately 13,100 km^2, with the main river lying about 248 km long occupying more than 85% of the Kelantan state. The estimated quantity of the annual rainfall in the basin is about 2383±120 mm, a large amount of which occurs during the Northeast monsoon between mid-October and mid-January. Kelantan River is the major river in the Peninsular State of Kelantan located in the Northeastern part of Malaysia. The river originated at the convergence of Galas and Lebir rivers close to Kuala Krai where it meanders near the coastal plain finally meeting South China Sea, about 12 km north of Kota Bharu the capital city of Kelantan state (Figure 1). Kelantan River's main reach possess larger tributaries that are located downstream. Furthermore, Galas and Lebir are known to also possess many tributaries escalated into the forested mountains of the country and are known to provide majority of the flow into the main Kelantan River. River Galas has two major tributaries viz., Pergau and Nenggiri that contribute to about 8000 km^2 or 54% from the total surface area of Kelantan's catchment (approximately 13,100 km^2). Figure 1 shows map of the study area with water quality monitoring sites.

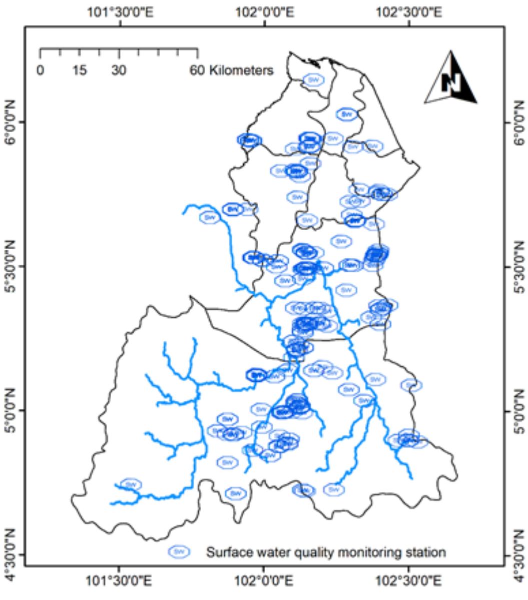


Figure 1. Map of the study showing sampling locations

Water Quality Data

Water quality data was obtained based on spatio-temporal availability from the Department of Environment, (DOE) Malaysia from 2005-2014. Fifteen water quality parameters that included total suspended solids (SS), pH, ammonium nitrate (NH_4-N), temperature (TEMP), conductivity (COND), turbidity (TUR), total dissolve solutes (DS), total solids (TS), nitrate (NO_3), chloride (Cl), phosphate (PO_4), calcium (Ca), potassium (K), magnesium (Mg) and sodium (Na) from several water quality monitoring sites were chosen for this study (Figure 1).

Statistical Analyses

Statistical analyses were conducted using multivariate statics such as Pearson correlation; PCA and CA in assessing water quality and establishing relationship between water parameters.

Principal Component Analysis (PCA). It is advisable to always determine the Kaiser-Mayer-Olkin (KMO) before performing PCA. This will help to assess how suitable it is to perform PCA on the intended data and to determine the sufficiency of the samples. The analysis can proceed if KMO value is found to be ≥ 0.5. In the current study; the value of KMO value was found to be 0.719. The PCA is the most commonly utilized multivariate statistical analysis. It functions by extracting variables into groups called principal components (PCs). These groups (PCs) are extracted along with their eigenvalues, variability (%) and cumulative values (%) (Ranjan et al., 2013). PCA was used in detecting patterns in data, which were presented based on their resemblances and dissimilarities. Pattern identification in a complex data is a difficult task, hence, the use of PCA is such scenario helps in providing good assessment (Smith, 2002).

Cluster Analysis (CA). This is a type of multivariate analysis that categorizes data into clusters according to their resemblance to each other and difference to other groups. CA is performed without making assumptions on the intended data so that structures or patterns can be discovered on the original data set (Mohapatra et al., 2011).

RESULTS AND DISCUSSIONS

Statistical Summary

The statistical summary of water quality parameters from 2005-2015 is shown in Table 1. The pH of the water in rivers of Kelantan showed both acidic and alkaline values (5.18 and 8.73 respectively). However, high SS (3380 mg L^{-1}) and high TS (3397 mg L^{-1}) might be because of high anthropogenic activities on the rivers. Urbanization around the river banks might influence the chemical reactions in the area due to leaching caused by sewage release (Abdulkareem et al., 2018). High TDS values ranging from 8.25 to 204 mg L^{-1} could also be attributed to anthropogenic activities such as agriculture. The order of availability of anions in the watershed is Cl>NO_3>PO_4 while that of cations were Na>Ca>Mg>K. There was high sedimentation in the area, which was evident from high values of turbidity (2780.00 NTU) in the basin.

Pearson Correlation

Guildford rule of thumb (Guilford & Fuchter, 1965) as presented in Table 2 was used in interpreting the relationship between water quality parameters. Results of the Pearson correlation matrix that describe the relationship between water quality parameters (Mor et al., 2009) are presented in Table 3. The results showed that most of the parameters were not correlated with one another even though high correlation existed in some few cases such as Turbidity-SS ($r = 0.856$, $p < 0.01$), DS-TS ($r = 0.982$, $p < 0.01$), DS-conductivity ($r = 0.82$, $p < 0.01$), DS-Turbidity ($r = 0.85$, $p < 0.01$), Na-Cl ($r = 0.76$, $p < 0.01$). While a moderate correlation was observed in the following parameters; Ca-conductivity ($r = 0.61$, $p < 0.05$), Ca-TDS ($r = 0.66$, $p < 0.05$), Mg-DS ($r = 0.52$, $p < 0.05$) and K-Mg ($r = 0.54$, $p < 0.05$).

Table 1

Summary statistics of water quality parameters in Kelantan River from 2005-2014

Parameters	**Unit**	**Minimum**	**Maximum**	**Mean**	**Std. deviation**
TSS	mg L^{-1}	1.00	3380.00	132.78	247.99
pH	-	5.18	8.73	7.11	0.55
NH_4-N	mg L^{-1}	0.01	2.60	0.07	0.17
Temperature	^{0}C	18.72	34.30	26.50	2.08
Conductivity	µS cm^{-1}	14.00	440.00	58.29	32.31
Turbidity	NTU	0.00	2780.00	155.09	259.47
TDS	mg L^{-1}	8.25	204.00	29.04	14.94
TS	mg L^{-1}	0.00	3397.00	158.12	246.10
NO_3	mg L^{-1}	0.01	1.72	0.19	0.20

Table 1 *(Continued)*

Parameters	Unit	Minimum	Maximum	Mean	Std. deviation
Cl	mg L^{-1}	0.52	74.00	2.91	4.02
PO_4-P	mg L^{-1}	0.01	1.00	0.03	0.06
Ca	mg L^{-1}	0.10	35.46	4.89	4.15
K	mg L^{-1}	0.10	11.05	1.34	0.82
Mg	mg L^{-1}	0.10	18.66	1.16	1.14
Na	mg L^{-1}	0.10	39.08	4.02	2.45

Ions such as Ca and Mg are the most abundant in rivers (Wollast & Mackenzie, 1983) while presence of Na and Cl with high correlation between them indicates the intrusion of saline water into rivers (Panteleit et al., 2001). SS is a major water quality parameter whose significance lies not only on the physical (light penetration) and ecological productivity (Parkhill & Gulliver, 2002; Rügner et al., 2013) but also is an indicator of other parameters such as phosphorus that are transferred on the surfaces of suspended sediments (Jones et al., 2011; Rügner et al., 2013). SS in the studied basin are a result of frequent flooding of the studied catchment.

Table 2

Guildford rule of thumb for interpreting correlation coefficient (Guilford & Fuchter, 1965)

r-value	Interpretation
0.0 to 0.29	Negligible or little correlation
0.3 to 0.49	Low correlation
0.5 to 0.69	Moderate or marked correlation
0.7 to 0.89	High correlation
0.9 to 1.00	Very high correlation

High correlation obtained between SS and turbidity in this study is attributed to the presence of suspended particles in rivers that causes high turbidity of the water by scattering light. The relationship between SS and TS and that of TS and turbidity were also found to be highly correlated in this study. These relationships are however, dependent on size, shape and type particles (Rügner et al., 2013). In Kelantan River basin where high agricultural and other anthropogenic activities are common along the riverbanks and where flow conditions in rivers are high during floods, more coarse particles may be suspended leaving smaller particles submerged. The moderate correlation between DS and Ca, DS and Mg and Mg and K can be interpreted as the mineral component of the river. High correlation between Na and Cl is an indication of seawater intrusion in northern part of Kelantan that is closed to South China Sea.

Table 3

Correlation matrix of water quality parameters in Kelantan River basin

Parameters	TSS	pH	NH_4-N	Temp	Cond	Tur	TDS	TSS	NO_3	Cl	PO_4-P	Ca	K	Mg	Na
TSS	1														
pH	-0.027	1													
NH_4-N	0.083	0.016	1												
Temperature	0.003	0.050	0.040	1											
Conductivity	-0.036	0.103	-0.004	0.12	1										
Turbidity	0.856**	-0.051	0.069	-0.01	-0.03	1									
TDS	-0.049	0.025	0.014	0.14	0.82**	-0.02	1								
TS	0.982**	-0.028	0.080	0.01	0.01	0.85**	0.01	1							
NO_3	0.264	-0.025	0.292	-0.01	0.10	0.25	0.08	0.249	1						
Cl	-0.016	-0.097	0.034	0.07	0.27	0.00	0.39	0.014	0.068	1					
PO_4-P	-0.037	0.051	-0.020	0.01	0.03	-0.01	0.01	-0.034	-0.076	0.060	1				
Ca	-0.013	0.091	-0.033	0.11	0.61*	-0.03	0.66*	0.032	0.119	0.128	-0.016	1			
K	0.114	-0.031	0.087	0.21	0.20	0.15	0.28	0.130	0.257	0.357	0.024	0.113	1		
Mg	-0.024	-0.017	0.022	0.09	0.46	-0.01	0.52*	0.013	0.114	0.335	0.009	0.482	0.54*	1	
Na	-0.038	-0.050	0.002	0.16	0.40	-0.04	0.49	-0.018	0.078	0.76**	0.046	0.194	0.494	0.483	1

*Notes: * significant correlation at $p < 0.05$ (2 tailed), ** highly significant correlation at $p < 0.01$ (2 tailed)*

Extraction of Components

Fifteen water quality parameters were subjected to PCA where 15 PCs (Table 4-5 and Figure 2) were identified but only 5 PCs were considered most useful because of their eigenvalue that is >1 (Arslan, 2013; Pathak & Limaye, 2011). PCs with higher eigenvalue contribute greater in controlling water quality (Abdi, 2003). In order to infer factors that are of utmost importance without altering the variance, varimax factor rotation was utilized (Kaiser, 1960). The eigenvalue-one criterion otherwise known as Kaiser criterion suggests that only PCs with eigenvalue value >1 should be adopted and interpreted. This is on the account that every one of the detected variables is responsible for one unit of variance to the total variation in the data.

Table 4

Eigenvalue, variability and cumulative % of extracted factors

Factors	Eigenvalue	Variability %	Cumulative %
PC1	3.70	24.67	24.67
PC2	2.95	19.64	44.31
PC3	1.49	9.94	54.25
PC4	1.21	8.07	62.32
PC5	1.07	7.16	69.48
PC6	0.98	6.52	76.00
PC7	0.87	5.80	81.80
PC8	0.81	5.42	87.22
PC9	0.64	4.29	91.50
PC10	0.42	2.79	94.30
PC11	0.29	1.96	96.26
PC12	0.21	1.41	97.66
PC13	0.18	1.21	98.88
PC14	0.15	1.02	99.90
PC15	0.02	0.10	100.00

Therefore, PCs with an eigenvalue more noteworthy than 1.00 are considered to contribute more variation than those with eigenvalue <1. PCA is conducted with the goal of reducing the number of experimental data into smaller components without damaging the actual meaning of the data sets involved. As such keeping PCs with less variance (eigenvalue <1) at the expense of those with more variance (eigenvalue >1) will defeat the aim at which PCA is conducted (O'Rourke et al., 2005).

Table 5
Factor loadings after Varimax rotation

Parameters	PC1	PC2	PC3	PC4	PC5
SS	0.07	0.96	0.08	-0.14	0.02
pH	0.01	-0.07	0.36	0.12	0.62
NH_4-N	0.07	0.18	-0.15	0.75	0.13
Temperature	0.24	-0.02	-0.06	0.05	0.55
Conductivity	0.76	-0.13	0.43	-0.02	-0.04
Turbidity	0.08	0.91	0.05	-0.16	0.01
DS	0.84	-0.13	0.32	-0.04	-0.09
TS	0.12	0.95	0.10	-0.16	0.01
NO_3	0.24	0.39	-0.08	0.62	-0.09
Cl	0.62	-0.06	-0.52	-0.20	-0.05
PO_4-P	0.03	-0.06	-0.08	-0.29	0.57
Ca	0.64	-0.10	0.58	0.05	-0.09
K	0.59	0.15	-0.44	0.12	0.15
Mg	0.75	-0.07	-0.03	0.03	-0.06
Na	0.75	-0.10	-0.46	-0.16	0.01
Eigenvalue	3.70	2.95	1.49	1.21	1.07
Variability %	24.67	19.64	9.94	8.07	7.16
Cumulative %	24.67	44.31	54.25	62.32	69.48

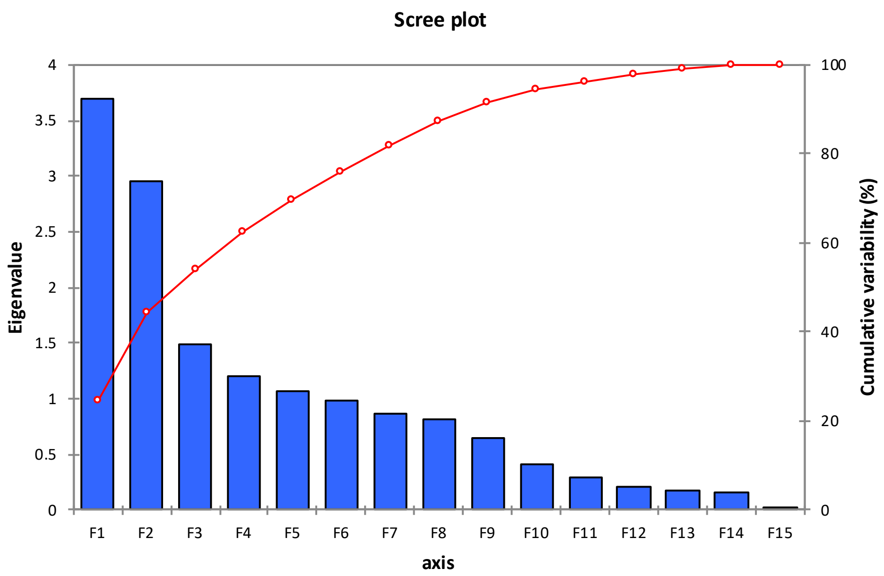


Figure 2. Scree plot of extracted factors (factors with Eigenvalue ≥ 1 are considered the most significant)

From the results of PCA (Table 4-5 and Figure 2) five PCs were identified which account for a cumulative value of 69% in the water quality. PC1 with eigenvalue of 3.70 recorded 24.67% of the total variance making it the PC controlling the most important processes influencing water quality (Yidana et al., 2010). In this PC, all the water quality parameters were observed even though some of the parameters recorded very little loading. A strong positive loading between electrical conductivity, TDS, Mg and Na was observed signifying that these parameters were dependent upon the contribution of other parameters. Ca was observed to be lower than both Na, which might be due to cation exchange activity that occurred naturally when ions of higher charges were replaced by those with lower charges ($Ca2+$ is replaced by $Na+$ at the exchange site). PCA is of utmost importance in hydrochemical analyses due to its ability to give inferences according to specific or multiple hydrochemical processes (Suk & Lee, 1999). A single PC can have one or more processes (Arslan, 2013; Kumaresan & Riyazuddin, 2008; Yidana et al., 2010) as observed in PC1 (Table 5) where high loading of DS and electrical conductivity resulting from seawater intrusion. Other processes such as Ca and Mg loading are because of weathering process (Kumaresan & Riyazuddin, 2008) although seawater intrusion may have dominated the whole scenario.

PC2 contribute 19.64% of the variability and it comprises a strong loading of parameters like SS, turbidity and TS (Table 5). The strong loading of SS, turbidity and TS may be because of influence of human activities along the riverbanks of Kelantan River basin making the water to attract more microorganisms. While the negative pH loading observed may be caused by high runoff activities in the watershed and leaching of sewage. An eigenvalue of 1.49 was observed in PC3 (Table 5) with a total variability of 9.94%. This PC shows a positive weak loading of pH, electrical conductivity and DS while a positive moderate loading of Ca and a moderate negative of Cl were observed. The positive moderate loading of Ca can be explained by weathering process occurring in the watershed. The weak loading of DS is an indication of small amounts of organic matter content in the water. PC4 has a total variability of 8.07% and eigenvalue of 1.21 (Table 5). A strong loading of $NH4-N$ and a moderate loading of $NO3$ were observed in this PC. This strong and moderate loading of nitrogen containing compounds is a clear indication of pollution caused by anthropogenic activities resulting from agricultural activities, sewage disposal and other domestic activities, which lead to build up of microorganisms (Sundaray, 2010). PC4 was reported to have a strong loading of pH and a moderate loading of temperature and $PO4-P$. The moderate loading of temperature is an indication of natural weather where high ambient air temperatures directly affect river temperatures. High loading of pH is an indication of photosynthetic activities caused by microorganisms such as algae and other aquatic plants United States Environmental Protection Agency (US EPA, 2012).

Cluster Analysis

When cluster analysis was performed on water quality parameters in Kelatan River basin, three major clusters were identified according to similarities existing between parameters and dissimilarities to other groups as outlined in Figure 3. Cluster 1 consisted of suspended solids, total solid and turbidity. The activities of microorganisms are directly related to turbidity (Mann et al., 2007). Higher activities of microorganisms in the river basin leads to oxygen depletion, which in turn affects dissolved oxygen concentration. In cluster 2, water quality parameters such as NH_4-N, NO_3, PO4-P, pH and temperature were observed. This cluster describes diverse processes involved in controlling the water quality of the basin. The presence of both NH_4-N and NO_3 is an indication of high level of anthropogenic activities taken place along the riverbanks in the watershed (Sundaray, 2010). In addition, pH and temperature are the major parameters controlling physiochemical and biological reactions taken place in water (Delpla et al., 2009; Nelson, 2000). Cluster 3 comprises electrical conductivity, dissolved solid, K, Na, Ca, Mg and Cl. This group gives an inference on the chemical activities in the watershed. The presence of Mg is an indication of freshwater recharge (Thilagavathi et al., 2012). While occurrence of Ca and Na signifies weathering processes taken in the river water while Cl may have been influenced by seawater intrusion as reported by Chidambaram et al. (2013).

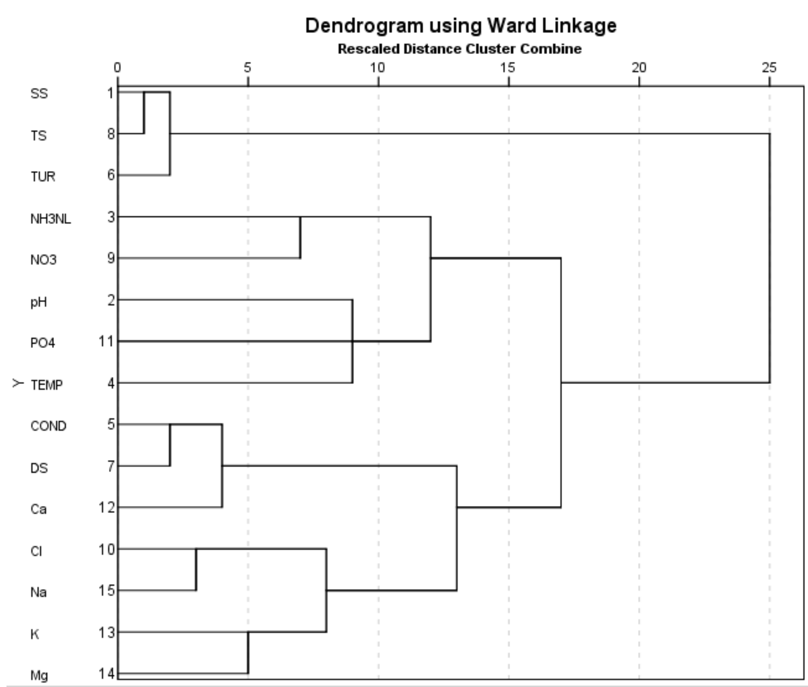


Figure 3. Dendrogram based on the clustering of water quality data in Kelantan

CONCLUSIONS

Fifteen water quality parameters from Kelantan River basin were used to conduct PCA and CA for identifying the parameters that influence water quality in the watershed. Water quality in the watershed is controlled by factors such as anthropogenic pollution, weather factors, seawater intrusion, weathering process and redox potential. PC1 was observed to be the most important component that is in control of water quality of the basin. Weathering processes and seawater intrusion are some of the processes that makes up the PC. Anthropogenic activities caused by high loadings of turbidity, suspended solids, total solids, NH_4-N and NO_3 are some of the processes influencing PC2 and PC4. Weathering activities also dominate PC3 due to moderate loading of Ca. High loading of pH and moderate loading of PO4-P suggesting that the PC is controlled by redox potential characterize PC5. Results of CA revealed that three clusters were observed. The first cluster comprises suspended solids, total solid and turbidity leading to the buildup of microbial activities. Cluster 2 is in control of diverse processes regulating water quality of the basin due to the presence of NH_4-N, NO_3, PO4-P, pH and temperature. The chemical activities involved in the watershed due to the presence of electrical conductivity, dissolved solid, K, Na, Ca, Mg and Cl are described by cluster 3. Results of this study will add to knowledge the specific water quality problems affecting the watershed, which can be utilized by land use managers and policy makers alike for future control planning.

ACKNOWLEDGEMENT

The authors wish to thank Department of Environment (DOE), Malaysia for providing the water quality data.

REFERENCES

Abdi, H. (2003). Factor rotations in factor analyses. In M. Lewis-Beck, A. Bryman, & T. Futing (Eds.), *Encyclopedia for research methods for the social sciences* (pp. 1-8). Thousand Oaks, California: Sage.

Abdulkareem, J. H., Pradhan, B., Sulaiman, W. N. A., & Jamil, N. R. (2017). *Prediction of spatial soil loss impacted by long-term land-use/land-cover change in a tropical watershed.* Retrieved March 1, 2018, https://doi.org/10.1016/j.gsf.2017.10.010

Abdulkareem, J. H., Sulaiman, W. N. A., Pradhan, B., & Jamil, N. R. (2018a). Long-term hydrologic impact assessment of non-point source pollution measured through Land Use/Land Cover (LULC) changes in a tropical complex catchment. *Earth Systems and Environment, 2018,* 1-18.

Abdulkareem, J. H., Pradhan, B., Sulaiman, W. N. A., & Jamil, N. R. (2018b). Quantification of Runoff as Influenced by Morphometric Characteristics in a Rural Complex Catchment. *Earth Systems and Environment, 2*(1), 145-162.

Adnan, N. A., & Atkinson, P. M. (2011). Exploring the impact of climate and land use changes on streamflow trends in a monsoon catchment. *International Journal of Climatology, 31*(6), 815-831.

Arslan, H. (2013). Application of multivariate statistical techniques in the assessment of groundwater quality in seawater intrusion area in Bafra Plain, Turkey. *Environmental Monitoring and Assessment, 185*(3), 2439-2452.

Chidambaram, S., Anandhan, P., Prasanna, M. V., Srinivasamoorthy, K., & Vasanthavigar, M. (2013). Major ion chemistry and identification of hydrogeochemical processes controlling groundwater in and around Neyveli Lignite Mines, Tamil Nadu, South India. *Arabian Journal of Geosciences, 6*(9), 3451-3467.

Dalakoti, H., Mishra, S., Chaudhary, M., & Singal, S. K. (2018). Appraisal of Water Quality in the Lakes of Nainital District through Numerical Indices and Multivariate Statistics, India. *International Journal of River Basin Management, 16*(2), 219-229.

Delpla, I., Jung, A.-V., Baures, E., Clement, M., & Thomas, O. (2009). Impacts of climate change on surface water quality in relation to drinking water production. *Environment International, 35*(8), 1225-1233.

United States Environmental Protection Agency. (2012). *Total alkalinity in water: Monitoring and assessment.* Retrieved April 26, 2018, from http://water.epa.gov/type/rsl/monitoring/vms510.cfm.

Guilford, J. P., & Fuchter, B. (1965). *Fundamental statistics in psychology and education.* New York: McGraw-Hill.

Jalali, M. (2010). Application of multivariate analysis to study water chemistry of groundwater in a semi-arid aquifer, Malayer, western Iran. *Desalination and Water Treatment, 19*(1-3), 307-317.

Jamaliah, J. (2007). Emerging trends of urbanization in Malaysia. *Journal of the Department of Statistics, Malaysia, 1*(2007), 43-53.

Jones, A. S., Stevens, D. K., Horsburgh, J. S., & Mesner, N. O. (2011). Surrogate measures for providing high frequency estimates of total suspended solids and total phosphorus concentrations. *JAWRA Journal of the American Water Resources Association, 47*(2), 239-253.

Jones, A. S., Stevens, D. K., Horsburgh, J. S., & Mesner, N. O. (2011). Surrogate measures for providing high frequency estimates of total suspended solids and total phosphorus concentrations. *JAWRA Journal of the American Water Resources Association, 47*(2), 239-253.

Juahir, H., Zain, S. M., Yusoff, M. K., Hanidza, T. I. T., Armi, A. S. M., Toriman, M. E., & Mokhtar, M. (2011). Spatial water quality assessment of Langat River Basin (Malaysia) using environmetric techniques. *Environmental Monitoring and Assessment, 173*(1-4), 625-641.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement, 20*(1), 141-151.

Kumaresan, M., & Riyazuddin, P. (2008). Factor analysis and linear regression model (LRM) of metal speciation and physico-chemical characters of groundwater samples. *Environmental Monitoring and Assessment, 138*(1-3), 65-79.

Kundu, R., Aura, C. M., Nyamweya, C., Agembe, S., Sitoki, L., Lung'ayia, H. B. O., … Werimo, K. (2017). Changes in pollution indicators in Lake Victoria, Kenya and their implications for lake and catchment management. *Lakes & Reservoirs: Research & Management, 22*(3), 199-214.

Mann, A. G., Tam, C. C., Higgins, C. D., & Rodrigues, L. C. (2007). The association between drinking water turbidity and gastrointestinal illness: A systematic review. *BMC Public Health, 7*(1), 256-262.

Mohapatra, P. K., Vijay, R., Pujari, P. R., Sundaray, S. K., & Mohanty, B. P. (2011). Determination of processes affecting groundwater quality in the coastal aquifer beneath Puri city, India: A multivariate statistical approach. *Water Science and Technology, 64*(4), 809-817.

Mor, S., Singh, S., Yadav, P., Rani, V., Rani, P., Sheoran, M., … Ravindra, K. (2009). Appraisal of salinity and fluoride in a semi-arid region of India using statistical and multivariate techniques. *Environmental Geochemistry and Health, 31*(6), 643-655.

Nelson, D. (2002). *Natural variations in the composition of groundwater; drinking water program.* Springfield: Oregon Department of Human Services.

O'Rourke, N., Hatcher, L., & Stepanski, E. J. (2005). *A step-by-step approach to using SAS for univariate & multivariate statistics* (2nd ed.). North Carolina: SAS Publishing.

Panteleit, B., Kessels, W., Kantor, W., & Schulz, H. D. (2001). Geochemical characteristics of salinization-zones in the coastal aquifer test field (CAT-Field) in North-Germany. In *Proceeding of 5th International Conference on Saltwater Intrusion and Coastal Aquifers-Monitoring, Modeling, and Management* (pp. 23-25). Essaouira, Rabat Morocco.

Parkhill, K. L., & Gulliver, J. S. (2002). Effect of inorganic sediment on whole-stream productivity. *Hydrobiologia, 472*(1), 5-17.

Pathak, H., & Limaye, S. N. (2011). Study of seasonal variation in groundwater quality of sagar city (India) by principal component analysis. *Journal of Chemistry, 8*(4), 2000-2009.

Ranjan, R. K., Ramanathan, A. L., Parthasarathy, P., & Kumar, A. (2013). Hydrochemical characteristics of groundwater in the plains of Phalgu River in Gaya, Bihar, India. *Arabian Journal of Geosciences, 6*(9), 3257-3267.

Rügner, H., Schwientek, M., Beckingham, B., Kuch, B., & Grathwohl, P. (2013). Turbidity as a proxy for total suspended solids (TSS) and particle facilitated pollutant transport in catchments. *Environmental Earth Sciences, 69*(2), 373-380.

Saiful, M., Iskandar, S., Azid, A., Juahir, H., Shakir, A., Saudi, M., … Amar, M. (2017). Control limit detection for source apportionment in Perlis River Basin Malaysia. *Malaysian Journal of Fundamental and Applied Sciences, 13*(3), 294-303.

Simeonov, V., Stratis, J. A., Samara, C., Zachariadis, G., Voutsa, D., Anthemidis, A., … Kouimtzis, T. (2003). Assessment of the surface water quality in Northern Greece. *Water Research, 37*(17), 4119-4124.

Smith, L. I. (2002). *A tutorial on principal components analysis introduction.* Dunedin, New Zealand: University of Otago.

Suk, H., & Lee, K. (1999). Characterization of a ground water hydrochemical system through multivariate analysis: Clustering into ground water zones. *Groundwater, 37*(3), 358-366.

Sundaray, S. K. (2010). Application of multivariate statistical techniques in hydrogeochemical studies - A case study: Brahmani-Koel River (India). *Environmental Monitoring and Assessment, 164*(1-4), 297-310.

Thilagavathi, R., Chidambaram, S., Prasanna, M. V, Thivya, C., & Singaraja, C. (2012). A study on groundwater geochemistry and water quality in layered aquifers system of Pondicherry region, southeast India. *Applied Water Science, 2*(4), 253-269.

Vieira, J. S., Pires, J. C. M., Martins, F. G., Vilar, V. J. P., Boaventura, R. A. R., & Botelho, C. M. S. (2012). Surface water quality assessment of lis river using multivariate statistical methods. *Water, Air, & Soil Pollution, 223*(9), 5549-5561.

Wan, I. (1996). Urban growth determinants for the state of Kelantan for the state's policy makers. *Buletin Ukur, 7*(1996), 176-189.

Wollast, R. & Mackenzie, F. T. (1983). The global cycle of silica. In S. R. Aston (Ed.), *Silicon geochemistry and biogeochemistry* (pp. 39-76). London: Academic Press Inc.

Yidana, S. M., Banoeng-Yakubo, B., & Akabzaa, T. M. (2010). Analysis of groundwater quality using multivariate and spatial analyses in the Keta basin, Ghana. *Journal of African Earth Sciences, 58*(2), 220-234.

Yu, D., Shi, P., Liu, Y., & Xun, B. (2013). Detecting land use-water quality relationships from the viewpoint of ecological restoration in an urban area. *Ecological Engineering, 53*(2013), 205-216.

Pertanika J. Sci. & Technol. 27 (4): 1971 - 1990 (2019)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Case study

Monitoring Urban Green Space (UGS) Changes by Using High Resolution Aerial Imagery: A Case Study of Kuala Lumpur, Malaysia

Junainah Abu Kasim[1]*, Mohd Johari Mohd Yusof[1] and Helmi Zulhaidi Mohd Shafri[2]

[1]Department of Landscape Architecture, Faculty of Design and Architecture, Universiti Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia
[2]Department of Civil Engineering, Faculty of Engineering, Universiti Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia

ABSTRACT

Urban green space (UGS) in a city is the foundation of natural productivity in an urban structure. It is also known as a natural cooling device that plays a vital role in the city as an urban lung, discharging oxygen to reduce the city heat and as a wall against harmful air pollution. When urbanization happens, UGS, including the gazetted areas, is essentially converted into an artificial surface due to the population's demand for new development. Therefore, identifying its significance is a must and beneficial to explore. The purpose of this study is to identify the 10 years of UGS change patterns and analyze the UGS loss, particularly in the affected gazetted zone. The study used available aerial imagery data for 2002, 2012, and 2017, and database record of green space. The study had classified UGS by using the Support Vector Machine (SVM) algorithm. The training area was determined by visual interpretation and aided by a land use planning map as reference. The result validity was then determined by kappa coefficient value and producer accuracy. Overall, the study showed that the city had lost its UGS by about 88% and the total gain in built up area was 114%. The loss in UGS size in the city could be compared to a total of 2,843 units of football fields, transformed forever in just 10 years. The uncontrolled development and lack of advanced monitoring mechanism had negatively affected the planning structure of green space in KL. The implementation of advance technology as a new mitigation tool

ARTICLE INFO

Article history:
Received: 26 January 2019
Accepted: 25 June 2019
Published: 21 October 2019

E-mail addresses:
junainah_UTM@yahoo.com (Junainah Abu Kasim)
m_johari@upm.edu.my (Mohd Johari Mohd Yusof)
helmi@upm.edu.my (Helmi Zulhaidi Mohd Shafri)
* Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

to monitor green space loss in the city could provide a variety of enhanced information that could assist city planners and urban designers to defend decisions in protecting these valuable UGS.

*Keyword***:** Gazetted areas, GIS, high resolution aerial imagery, image segmentation, land cover change, urban green space (UGS)

INTRODUCTION

The importance of urban green space (UGS) is already known for decades. It is demonstrated by many collected reviews whereby most findings claimed that green space can enhance city ambience, improve urban dwellers' well-being and generate a positive monetary asset in a strategic way (Cilliers, 2015; De Ridder & Lefebre, 2003; Wolch et al., 2014). It is not only seen as a natural hardscape to provide a recreational activity space, but is also functioning beyond appraisal. Moreover, many studies have claimed that the presence of UGS could help a city to combat urban heat islands, and thus improve all city body elements, for instance, its heat and air quality, as well as promoting bodiversity growth (Liu & Shen, 2014; Tani & Chang, 2017). However, urbanization process has put more pressure to the city development; thus, it is reflected in the growing city size and change in land use arrangement (Nor et al., 2017). Therefore, some of the UGS are being transformed and converted into hard surface that is fully covered with concrete and asphalt, with no space left for any green space to live in (Jiang & Tian, 2010). Similarly, Malaysia as a developing country also experiences rapid urbanization, especially in Kuala Lumpur (KL) that has an acreage of 243 square km. The hasty development process has attracted population to reside in KL; thus, more land are acquired for new development to meet the urban dwellers' demands (Isa et al., 2017) . Generally, the uncontrolled new economic development, particularly in KL city center, has expanded the city size towards suburbs, and subsequently altering the natural resources (Elsayed, 2012).

Monitoring green spaces within an entire city is important in enhancing information which enable city planners to know the current events and regularly update the green space database. Most of the local councils including Dewan Bandaraya Kuala Lumpur (DBKL) use Geographic Information System (GIS) to track land use changes and as a support tool for controlling urban development. However, the available database is obsolete and outdated, and relyon the manual inventory information (Yusof, 2012). The use of high-resolution aerial imagery has great potential to help city planners in time to extract, update, and detect land changes by applying advance spatial techniques and tools (Noorollahi, 2005). With the capacity and ability to differentiate land cover types at a larger scale, remote sensing is widely used and explored to map and detect green space changes . Available datasets of the USA imaging satellite, such as GeoEye, QuickBird, IKONOS, and WorldView will give a real world imagery with ground accuracy distance up to 15km. It is also often used by many scholars to study environmental changes (Michel et al., 2015). However, the spatial detail of these datasets is not sufficient to capture, classify and segment the small patches of green space, and thus limits the spatial distribution and classification of green space present in the urban fabric form (Vatseva et al., 2016). Recently, many studies have used aerial photography to measure urban growth and land use changes due to greater accuracy and pixel resolution in the image covers and temporal data. For example, Moskal et al. (2011) used the 1m, 4-band infrared of aerial photography to monitor urban tree cover area in Seattle, USA. The higher spatial resolution used in this study generated a good classification with 79.7% overall accuracy. The significant strength of hyper spatial imagery was that it could map very small urban features, for instance, small patches of green space or shrub clusters. Besides, the limited spectral bands also allowed spatial and spectral comparability, and thus details in land changes detection can be attained (Crommelinck et al., 2016).

Land cover classification is widely used in many fields, such as urban planning, environmental change and land resources mapping. It digitally illustrates on how the land has been developed, conserved and changed through time. To classify the land into beneficial information, a good classifier should be determined. According to Bahari et al. (2014), there were three factors that should be considered when selecting the classification method, such as accuracy, processing speed and efficiency, to handle huge data input. For example, Abebe and Megento (2016) used the maximum likelihood classification (MLC) algorithm to classify land cover into 12 classes. The study had proven that the city of Addis Ababa had UGS loss with an annual rate of 5.9% over 30 years through high detection of changes from an aerial imagery source. However, classifying aerial imagery by using MLC needs a larger training area besides assuming the sample in use to be normally distributed. This can minimize the optimal separating image into various land use classes and affect the accuracy agreement. Beyond the conventional method classifier like MLC, Li and Shao (2014) presented the object-based approach in detecting land changes at Tippecanoe County in Indiana State, USA from 1 m resolution imagery with 5-feet DEM. The study used multi-threshold (MT) segmentation to classify the land use classes. The algorithm was capable to classify three categories of vegetation types, such as trees, grass and crop, through its own spatial distinct and spectral features. However, even though MT segmentation approach was effective in separating an image into meaningful object, its processing time was rather much longer than quadtree-based (QT) and SVM segmentation, and thus requires high hardware performance (Hsu et al, 2010). In recent years, there were efforts to develop and turn out the classification method to be more reliable and efficient in computation time and result accuracy. SVM is an artificial intelligence and novel machine-learning classification algorithm. For example, the method can apply minimal training area within a larger scale study to extract comprehensive information on land cover. SVM also requires no assumption on data type therefore less tedious (Rudrapal & Subhedar, 2015; Shi & Yang, 2015). Bahari et al. (2014) implemented SVM to classify tropical land cover in Klang Valley, Malaysia. The study successfully classified the high-resolution imagery into 10 classes without any pixels being unclassified in a quick time processing. Furthermore, the study showed that the overall accuracy for Klang Valley tropical land cover classification was 97.1% with kappa coefficient at 0.96. It indicated that the classifier can produce high agreement between sample areas and classified data. Therefore, this study was carried out by using SVM classifier to perform image classification due to its ability to handle the complexity of features and optimal assessment for separating the features into classes with maximum margin. The aim of this study is to present the distribution and analyze the UGS changes in KL with functional classification of SVM based on available high-resolution aerial imagery data and GIS application. Furthermore, the wide range of output generated from the multiple temporal aerial imagery can improve data reliability and develop high performance to interpret the visualized phenomenon in the real world.

STUDY AREA AND DATA SETS

Study Area Boundaries

The study area of Kuala Lumpur is the most populous city in Malaysia with an estimated population of 2.2 million by 2020. It covers an area of about 243km^2 (Figure 1). It is located at the west of Peninsular Malaysia at 3°8.472′ N, 101°41.1918′ E and originally formed at the confluence of Sungai Gombak and Sungai Klang (Dewan Bandaraya Kuala Lumpur, 2018). The city was selected as the ideal test-bed area for green space classification because in the last two decades, KL experienced rapid changes in commercial activities and the resided population into the city center (Elsayed, 2012). Furthermore, the intention to develop KL into a major Asia commercial hub has tremendously changed the green area patterns in this city (Rakhshandehrooet al., 2015). According to Yaakup et al. (2005) KL has boosted its built-up areas since 1950 and has an estimated green space loss of 50% due to the high demand for housing developments, high grade infrastructure and modern public facilities. Since then, KL has received a huge transformation plan in land use activities, including the building of Kuala Lumpur Convention Center (KLCC), and has continued to undergo new rapid mega-project developments. Therefore, it is beneficial to visualize these trends and employ advanced approach in order to analyze the environmental changes in the city.

A definition of UGS was specified as the first task in this study to determine the range of classification assessment based on the objective criteria. This study used the term 'urban green space' that contained the elements of greening area covered in urban setting, including cemeteries and institutional green space. It was comprehensive and covered almost every point in urban area, especially KL whereby reserved forests, undeveloped land, green

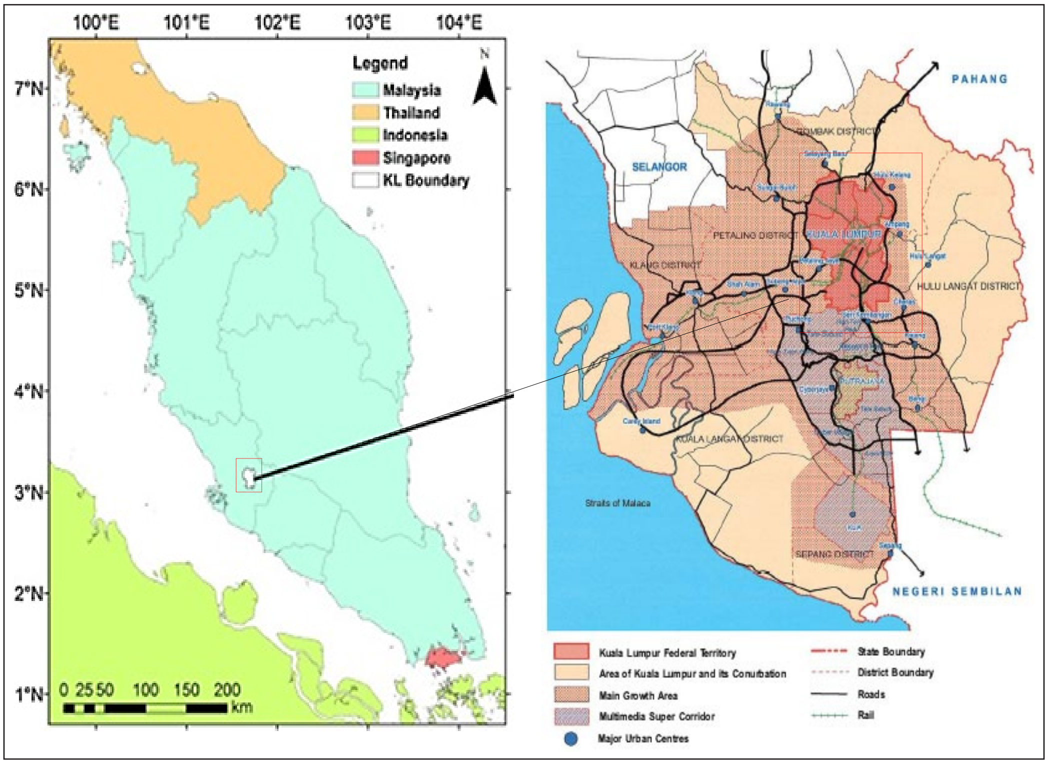


Figure 1. Map of Kuala Lumpur
Source: Dewan Bandaraya Kuala Lumpur, 2013

corridors, recreational areas, open land within the built-up areas, riverside, golf course and public amenity (institutional green areas and cemeteries) were included. This definition also includes water areas that provide greenways that enable people to connect with nature. However, this study did not capture the vertical green space and individual private green area in a housing gated scheme due to access limitation and data availability. Table 1 shows the full extent of green space classification adapted from Kuala Lumpur City Plan 2020 (KLCP 2020) (Dewan Bandaraya Kuala Lumpur, 2018) that were considered in this study.

Table 1
UGS classes for Kuala Lumpur

Main Class	Sub-Classes
Urban Green Space	District Park, City Park and Local Park
	Urban Greenery in Housing Neighborhood
	Playgrounds/Playing Lots/Playing Fields
	Urban Greenery in Public Facilities
	Greenery in Sport Facilities and amenities
	Forest Reserve
	Cemeteries

Table 1 *(continue)*

Main Class	Sub-Classes
Urban Green Space	Green Conservation Area (Hill, River, Lakes)
	Controlled Green Area (Public Open Space, Private Open Space, Transmission Line, Redevelopment Areas, Infrastructure Corridor)
	Open land within built-up areas

Data Sets

In this study, the orthophotography data for 2007, 2012 and 2017 were gathered from the Department of Survey and Mapping Malaysia (JUPEM). The city boundary maps, green area development data and spatial land use map were obtained from the GIS division of DBKL. Detailed information related to the data used is listed in Table 2.

Table 2
Detailed information on the data used

Types of Data	Description	Data Source
Orthophotography	Ground Sample Distance (GSD) = 4cm Scale = 1:15,000 Altitude = 60m/200ft Radiometric Resolution = 8 bit Digital Number Values (DN) = range 0-255 Format = Tagged Image File Format (TIFF)	JUPEM
Planning Map	Land use planning map (year 2007, 2012, 2017) = Shape file Green space database = existing and gazette information inventory KL administrative boundary = Shape file *All data used GDM 2000 reference (Geocentric Datum Malaysia 2000)	DBKL

METHODS

The overall process of green space segmentation and classification that was extracted from orthophoto data is illustrated in Figure 2. The segmentation process with selected parameters and accuracy assessment are described in the following sub-section. The software used for this analysis was ArcGIS Pro due to the capability and enhancement on this application to support the entire classification workflow, such as segmentation, training sample collection, classifier procedure, class merging and editing, and accuracy assessment in one comprehensive image analyst tool (Liedtke & Hu, 2018). The use of this application has received wide attention by many researchers and they claimed that the classification tools provided quick access to the main functionality with consideration of time consuming and the hardware capacity to handle big data processing (Kwak, 2016; Nakata-Osaki et al., 2018; Santos et al., 2016).

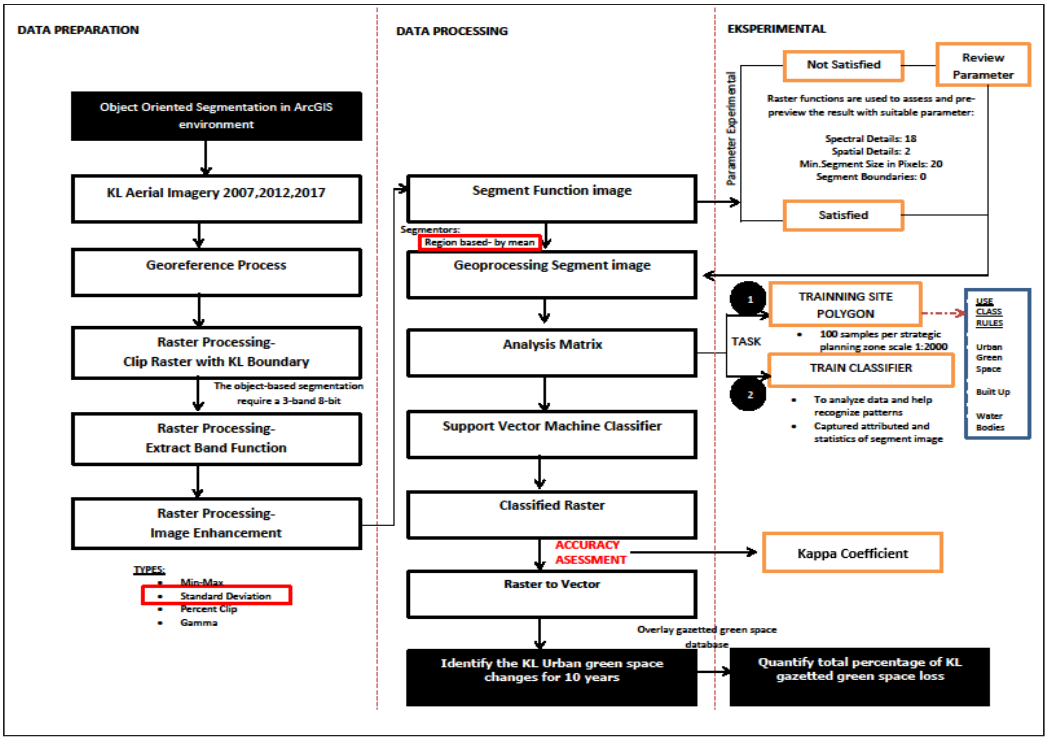


Figure 2. Flowchart of segmentation and classification methodology for quantifying the loss of gazetted green space in Kuala Lumpur by using ArcGIS Image Analyst

Image Pre-Processing

The imagery pre-processing and data management were the initial processes to be conducted. After reprojecting all datasets to the coordinate system of Geocentric Datum Malaysia 2000 (GDM 2000), the image enhancement analysis function was utilized to derive all imagery data having the same resolution conditional of 0.5m with 3-band (RGB) 8 bit-unsigned. The conditional helped to reduce spectral dimension of multivariance images. As a result, it could present three uncorrelated features that simply distinguish vegetation cover and impervious surface contained in RGB bands. Besides, the 0.5m higher resolution data were used to perform high-detection on-screen and recognize individual pixels so as to increase the chance to have the finest object segmentation and processing image classification successfully (Li & Shao, 2014).

Segmentation and Classification

For this study, the object-based segmentation technique was applied to identify KL green space changes in a 10- year time series. The technique comprises two steps. Firstly, the parameter needs to be experimented and tested by using the function of segment-mean shift. This tool allows the output to be generated temporarily and modifies according to the amount of spatial and spectral smoothing. The parameter input experimentation for spectral, spatial, and segment size per pixel could determine the smoothness, extract color, texture and location of compactness of features from each pixel (Amalisana et al., 2017). Secondly, after recording the suitable parameter that showed the best derivation of features or interests, the geo-processing of segment mean was shifted by using the raster function tools.

In the experiments, the selected segmentation parameters were chosen to balance between spectral and spatial details. The maximum valid value range was from 1 to 20. The spectral detail was dominating the evaluation in achieving the highest segmentation result. The highest value could obtain high homogeneity, and hence to a certain degree could increase the accuracy agreement. Besides, the use of higher spectral value would contribute to greater features discrimination, and thus classify the different land cover classes effectively. Meanwhile, the smaller value in spatial detail could enhance the imagery shape and characteristics of the pixels subsequently creating a smoother segmentation output. In addition, the determination of the maximum segment size parameter was related to the minimum unit of mapping (MMU) or the amount of pixel details. The less detail values will take longer in analysis processing times due to generating a smoothing image segmentation. In a summary, the concept of segmentation parameter sets is illustrated in Figure 3.

In this study the most suitable parameter was successfully identified at Stage 4 (Table 3). The selection of parameter was chosen by using recursive visualization (Figure 4) by the analyst, with the understanding of the segmentation parameter conceptual adapted from Liedtke and Hu (2018).

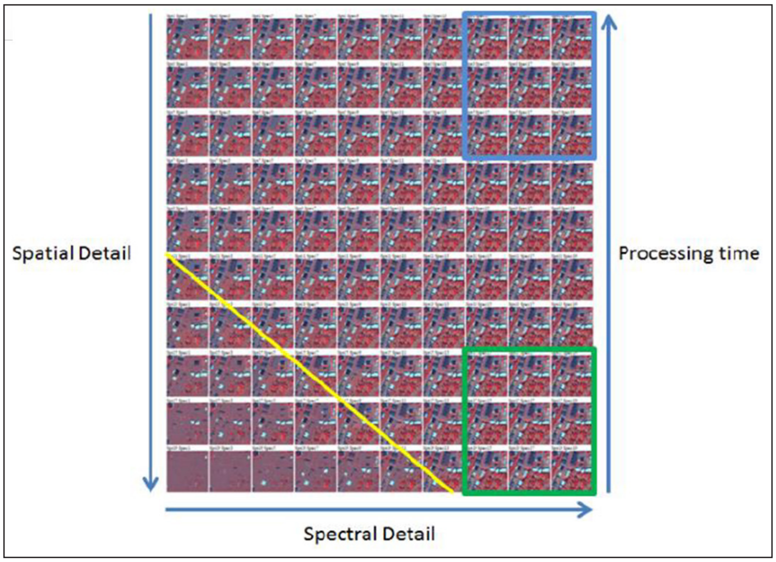


Figure 3. Segmentation parameter conceptual (*Source*: Liedtke & Hu, 2018)

Table 3
Algorithm parameters test and used

Number of Test	Spectral Detail (Max. value 20)	Spatial Detail (Max. value 20)	Min.Segment Size in Pixels (Max. value 20)
1	5	2	10
2	10	10	18
3	18	15	20
4	18	2	20

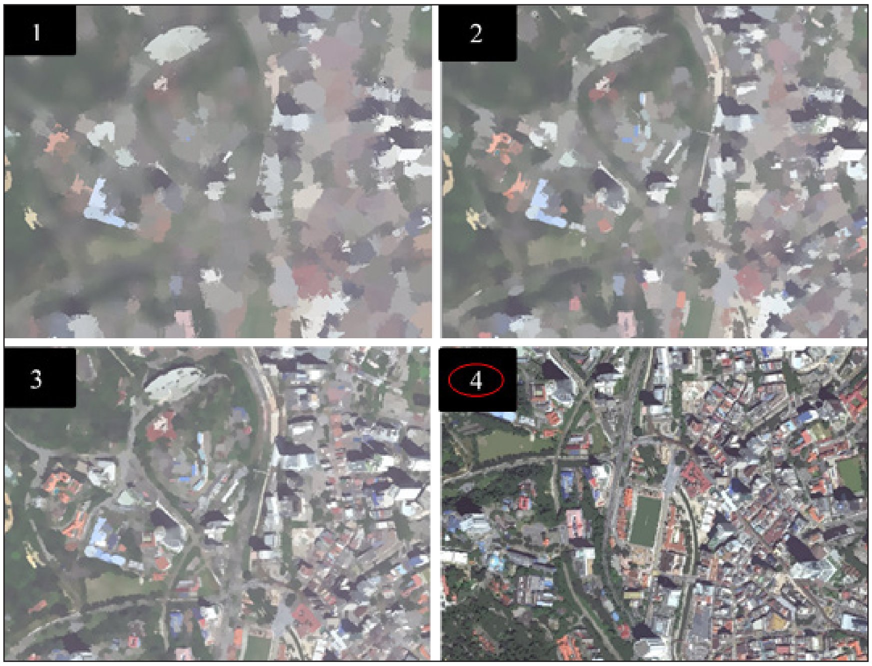


Figure 4. Image segmentation processing with different parameter testings through segment-mean function algorithm. Test 4 image indicated significant results.

Based on the selected parameter, the spectral details were made larger by 18/20 so that the finest object could be detected, especially for research conducted in urban areas. This was supported by choosing the lower spatial details to create smoother segmented images. According to Kaur and Kaur (2014) there is no specific technique as the best way to do segmentation because the parameter set is developed based on the subject matter to be studied and to suit with the image type. Besides, the requirement to obtain a satisfied complex parameter is crucial and difficult to find (Wężyk et al., 2016).

Image classification can be done by two algorithms that are supervised and unsupervised. The supervised classification was employed in this study and it required training samples that will help to identify the land cover classes. The reference data that included land use map and green space database were used to assist in supervising the training site selection according to the features of interest. To address the issue of time consumption in doing the sample sites, the image was divided into six zones for effectively managing the classification task. The division of six areas was digitized based on the KL strategic planning zones, as shown in Figure 5. Each zone represented its distinctive spatial pattern and land use activity in order to create more variation segmentations in terms of shape and texture. The zoning also considered the green space distribution patterns to avoid bias in certain dense areas.

This study had at least 200 training areas with a minimum of 10-30 polygons for each strategic zone. To ensure that the training samples were homogenously distributed, this preliminary classification was overlaid with the original image to determine its logical group. SVM classifier was used for classification assessment based on the input generated from the training samples. Besides, the classifier was considered as an easy procedure

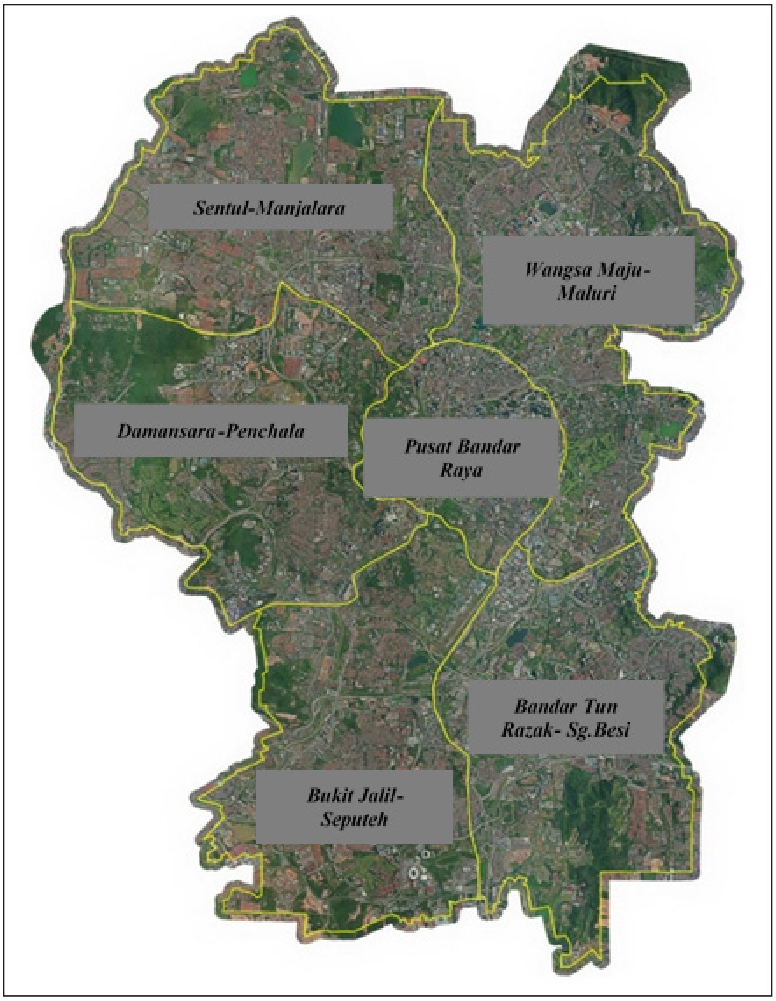
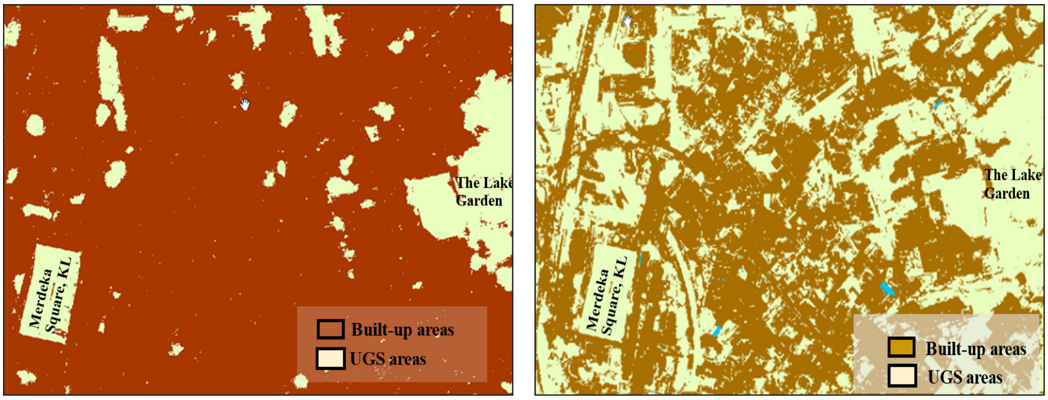


Figure 5. Kuala Lumpur Structure Plan (2004) with six KL administrative boundary zones for more strategic land use and planning distribution



(a) (b)

Figure 6. (a) Image representing the classification output using MLC technique and (b) using SVM technique

to operate because it did not require a large sample, and thus can address the issue of processing time. The SVM extraction of each object was more accurate and detailed as compared to the maximum likelihood (MLC) algorithm. Figure 6 shows the classification output by using SVM algorithm as compared to the MLC algorithm.

Generally, SVM is a non-parametric binary linear classifier. It works by identifying the optimal separating hyperplane (such as a dividing line to separate the data into two set) in a high space by using the kernel approach to locate optimal boundaries between classes (Rudrapal & Subhedar, 2015). It can produce a good classification with less amount of sample data and did not require the features selection in the pre-processing assessment because it can independently function based on fully dimensional hyperspectral data. The easier way to work with SVM classifier is by using linearly separable classes. Adapting from the work by Taati et al. (2015), the SVM algorithm is expressed as,

$$WXi + b \geq +1; \text{ which is for all } y = +1 \text{ (member of Class 1)} \quad [1]$$
$$WXi - b \leq -1; \text{ which is for all } y = -1 \text{ (member of Class 2)} \quad [2]$$

Based on Equation [1] and Equation [2], the numbers of training data are represented as $\{Xi, Yi\}$, i =1, 2....k, where X € Rn is denoted by n-dimensional space and Y € {+1, -1} is a class label. Then, all classes involved are considered as linear and separable if a vector of W is present and perpendicular to the linear hyperplane. For the two classes, Class 1 is represented as +1 while Class 2 as -1. The hyperplane is capable to discriminate the data point into respective classes. The kernel function in the SVM algorithm offers a solution to locate the training sample into one of the two categories, making the decision accurate and error free (Rudrapal & Subhedar, 2015). In additional, Figure 7 illustrates the basic concept on how the SVM functions. They are unlimited numbers of hyperplanes that will be involved in particular areas of study and SVM will consider selecting the hyperplane with maximum margin. Then, SVM will calculate the gap between the margin range, which is from the reference vector (classifier) and the training point (support vector). The main idea is to maximize the clearance distance for which the gap is large.

Therefore, in this study, after processing the image segmentation and classification procedure, there are three main land covers that were successfully identified, namely urban green space areas, built-up areas, and water bodies (Table 4).

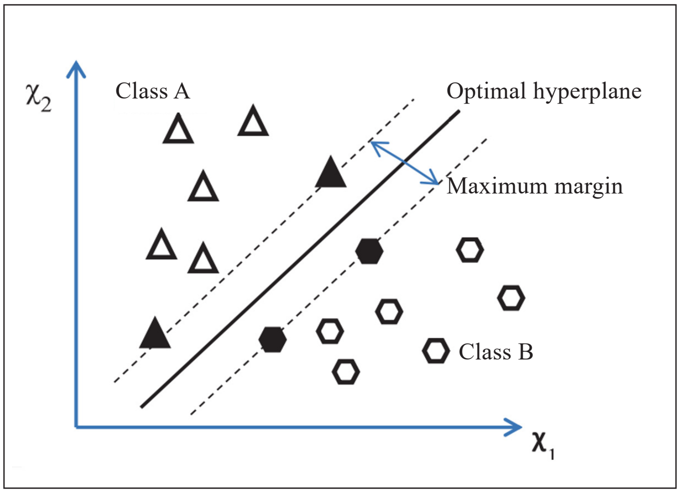


Figure 7. Conceptual of SVM functional (*Source*: Bahari et al., 2014)

Table 4
Identified classes by supervised classification

No.	Land classes	Description
1	Built-up	Residential, commercial, Industrial, Roads, Railway, mixed urban and developed lands
2	UGS	District Park, City Park and Local Park, Urban Greenery in Housing Neighborhood, Playgrounds/Playing Lots/Playing Fields, Greenery in Sport Facilities and amenities, Forest Reserve, Cemeteries, Controlled Green Area (Public Open Space, Private Open Space, Transmission Line, Redevelopment Areas, Infrastructure Corridor), Open land within built-up areas
3	Water Bodies	Lakes and ponds, wetlands

Accuracy Assessment

In this study, test sample areas (pixels) of approximately 50 objects or polygons for each class were selected by visual inspection through the entire images and with consideration the sample are well distributed. Then, error matrix (also known as confusion matrix) statistic was applied to test the degree of difference between the automated classified images with the sample areas. The validation of the output was important to determine the ratio of pixels identified for each category and to find out classification error, whether over-estimation or underestimation (Shubho et al., 2015). The result was presented in statistical form of producer accuracy and overall accuracy. According to Anderson et al. (1976) 85% is the minimum accuracy value which is acceptable and satisfactory to classify the images. In additional, kappa coefficient value was demonstrated as less than 1 or equivalent to 1. A value near or equal to 1 indicates that variables are in perfect agreement and significant to illustrate the real event.

RESULTS AND DISCUSSION

Classification Accuracy Assessment

Table 5 shows the summary of accuracy result for all three sets of the imagery data while Table 6 illustrates the detailed accuracy of the image classification for 2017 as a sample result. The overall accuracy indicated that the percentage of successful pixels was correctly classified accordingly to their land use classes with the application of SVM classifier. The results also demonstrated that most of the SVM classification and sample areas were in perfect agreement, which were the kappa coefficient for the three set images were recorded near to value 1.

Table 5
Accuracy assessment for supervised classification using SVM classifier

Year of Data (Orthophoto)	Overall Accuracy	Kappa Coefficient	Indicator	Significant Agreement
2007	91.2%	0.84	<0 no agreement 0 – 0.20 slight 0.21 – 0.40 fair 0.41 – 0.60 moderate 0.61 – 0.80 substantial 0.81 – 1 almost perfect agreement	Almost perfect agreement
2012	94.67%	0.89		
2017	93.3%	0.83		

Table 6
Error Matrix (%) comparing image classification for 2017 to the sample areas (pixels) data

	Sample areas (Pixels)					
	Classes	UGS	Built-up	Water Bodies	Total Classified Pixels	User Accuracy (%)
Classification Data (Pixels)	UGS	**40**	3	0	43	93.0
	Built-up	5	**68**	0	73	93.2
	Water Bodies	2	0	**32**	34	94.1
	Total Sample Area	60	80	160	**150**	
	Producer Accuracy (%)	66.7	85	20		
	Overall Accuracy		93.3			
	Kappa Coefficient		0.83			

Referring to Table 6, the automated image classification for 2017 gives 93.3%accuracy. Water bodies were well segmented, and thus recorded the highest accuracy value. UGS and built-up areas delivered slightly misclassified pixels whereby confusion occurred due to similar spectral values and homogenous in spatial form. However, both classes showed a higher accuracy of 93.0% and 93.2%, respectively. Therefore, the accuracy assessment was reliable and derived further significant output.

The Changes of Green Space from 2007 to 2017

The analysis of UGS changes based on the SVM classification method had revealed that in the first period (from 2007 to 2012), KL UGS area had slightly decreased by 191 ha and tremendously reduced by 1515 ha in 2017. Besides, it can be observed from Table 7 that the built-up area had greater increased in 2017 with 12,148 ha, which covered the study area by 50%. Furthermore, water bodies were extremely lost from 2007 to 2012, which was 117.16 ha, and then remained stagnant towards 2017. In summary, the total proportion of UGS loss from 2007 to 2017 was 88%, while a total gain of built up areas was 114%. The result is also visualized in Figure 8.

Table 7
The proportion difference gained and loss for each type of land use classes from 2007 to 2017

Land Use Class	2007		2012		Difference Gained / Loss from 2007 till 2012		2017		Difference Gained / Loss from 2007 till 2017	
	Ha	%	Ha	%	Ha	%	Ha	%	Ha	%
UGS	13,102	54	12,911	53	191	-99	11,396	47	1515	-88
BUA	10,325	42	10,633	44	308	103	12,148	50	1515	114
WB	873	4	755.84	3	117.16	-87	755.84	3	0	0
SUM	24,300	100	24,300	100			24,300	100		

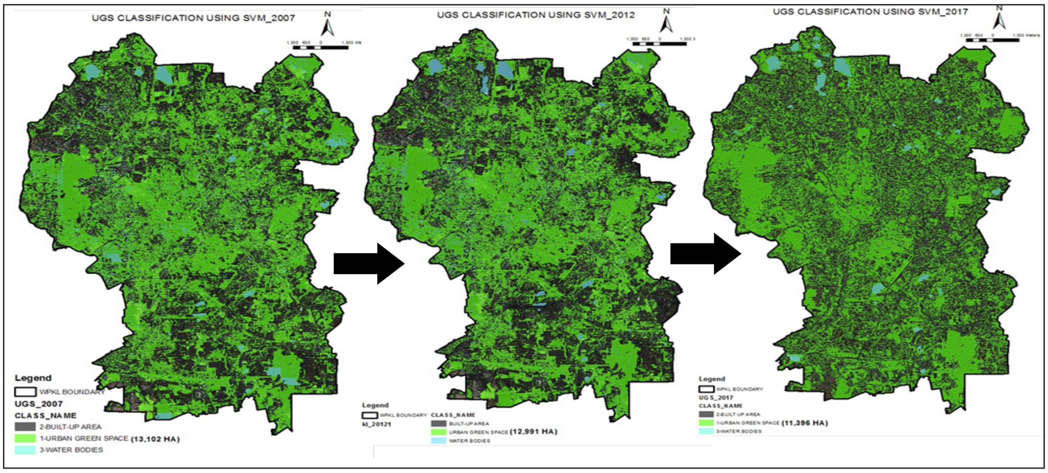


Figure 8. Changes in the green space in Kuala Lumpur from 2007 to 2017 based on SVM classification method

Generally, the size of UGS loss in KL can be compared to a total of 2,843 units of football fields that has transformed into built up areas. The result also demonstrated that the UGS in KL had struggled to survive in the last 10 years due to the demand of new development that had acquired more lands. Furthermore, this event was related to the intention of KL to become a world class city by 2020 with the provision of various mega developments, such as high grade infrastructure and modern public facilities (Nor et al., 2017; Yusof, 2012). The vision has directed to the uncontrolled land management and land scarcity in KL had pushed the new development to use available green areas.

The Loss of Green Space in the Gazetted Zone Areas

The trend for a 10-year period showed that the net green space loss was randomly located at the northeast and spreading out to the west and south of the city's boundary (Figure 9). Most of the destructed areas were located near or within the gazetted zones. A recent government report (Dewan Bandaraya Kuala Lumpur, 2018) examined the total green space that was legally declared as 'prohibited development areas' and recorded at about 978.42 ha. However, the figure can be argued because through the classification analysis, the potential green space that was identified to be preserved and protected was tremendously transformed into new developments. This can be proven by referring to Figure 9 that demonstrates four samples of the gazetted parks that were allowed to develop into a high-rise affordable housing project. These four sites obviously separated the object features into two classes which were built-up area and gazetted green space zone.

In addition, it can be observed from Table 8 that in Metropolitan Kepong Park, the gazetted zone has lost about 4.04 ha, which covers 4.8% of its total area. Furthermore, Taman Wahyu Park, Tasik Permaisuri Park and Bukit Jalil Technology Park was decreased by 14.49 ha, 3.23 ha and 1.63 ha, respectively.

Table 8
The percentage of green space loss in selected gazetted zone areas

Site	Acreage 2007-2012 (Hectares)	Acreage 2017 (Hectares)	Allocation for affordable housing land (Hectares)	% Loss
A	83.17	79.13	4.04	4.80
B	88.57	74.08	14.49	16.35
C	9.32	6.09	3.23	34.65
D	53.39	51.76	1.63	3.05

Overall, the gazetted green space loss in KL was 1,706 ha, which was 12% out of the total city areas. However, despite the facts showing a small loss, the relatively pint-sized of green spaces throughout the years will lead to higher air temperature (Chejarlaet al., 2016), can harm urban wildlife homes (Karuppannan et al., 2014) , was significant to

Figure 9. The net loss of green space in Kuala Lumpur and sample sites of four gazetted parks affected by development

increase water volume, and thus causing flood (Kim et al., 2016), erosion and polluted water problems (Mansor & Said, 2008).

In summary, the declining trend of the green space showed that urbanization had continuously occurred without a limit even though there active action plans were strategized, for example, green initiatives that aimed to plant 100,000 trees by 2020 (Government of Malaysia Performance Management and Delivery Unit, 2016). Besides, green space was tremendously sacrificed for the objective of political force and investment prophecy. New development had aggressively entered the gazetted zone to develop affordable housing projects, for instance, *Rumah Mampu Milik Wilayah Persekutuan* (RUMAWIP), *Perumahan Rakyat 1 Malaysia* (PR1MA) and *Perumahan Penjawat Awam 1 Malaysia* (PPA1M). It can be argued that with the limited land available in the city to cater for the population's demand, the green space is seen as the only space left to be developed. Obviously, the green space was now meant not only for environmental improvement but had to make way for social and economic development too. Therefore, a comprehensive UGS planning framework through various initiatives must be strategied in an effective way and should be aided by visualize monitoring system to ensure that the profits of UGS could dispense to all urban sector developments.

CONCLUSION

Overall, the classification assessment done in this study had revealed the amount of remaining urban green space in KL. The land cover change analysis revealed that the UGS had constant decrease over the last 10 years. In summary, the city had UGS loss of about 88% and the total gain of built up areas was 114%. The rapid development has grown the size of KL city, and thus altered the left-over green space. The integration between GIS and high-resolution aerial imagery can give high accuracy evidence and can play an important role in monitoring and controlling urban development. The presented case study showed the high performance of automatic extraction of green space from three sets of high-resolution aerial imagery data for the year 2007, 2012, and 2017. The accuracy result for all three sets of the imagery data were very high, as recorded at 91.2%, 94.67%, and 93.3%, respectively. Besides, the method applied was significant to map KL green space changes and prove that the event had more accuracy by using a SVM classifier. The outcome generated from this spatial analysis provided an up-to-date information that served genuine green space data, and thus improved the decision-making process. The outcome can also help the Kuala Lumpur City Hall to identify the location where natural preservation areas are threatened. Furthermore, this spatial form analysis would contribute for planning the protection of green space across KL's boundary and prioritizing specific regulation in dealing with this event.

ACKNOWLEDGEMENTS

The research team acknowledges the Planning Department of KL City Hall (JPRB, DBKL) and Department of Survey and Mapping Malaysia (JUPEM) for providing aerial imagery data and cadaster lot database. We are also thankful for the research grant, IPS Putra Grant Research (GP-IPS/2017/9586600) from Universiti Putra Malaysia and KL City Hall Scholarship for the research funding and assistance throughout this study.

REFERENCES

Abebe, M. T., & Megento, T. L. (2016). The city of Addis Ababa from 'Forest City' to 'Urban Heat Island': Assessment of urban green space dynamics. *Journal of Urban and Environmental Engineering*, *10*(2), 254–262.

Amalisana, B., & Hernina, R. (2017, September 27-28). Land cover analysis by using pixel-based and object-based image classification method in Bogor land cover analysis by using pixel-based and object-based image classification method in Bogor. In *The 5th Geoinformation Science Symposium 2017 (GSS 2017)* (pp. 1-7). Yogjakarta, Indonesia.

Anderson, J. R., Hardy, E. E., Roach, J. T., & Witmer, R. E. (1976). *A Land Use and Land Cover Classification System for Use with Remote Sensor Data* (Vol. 964). Washington, D.C.: US Government Printing Office.

Bahari, N. I. S., Ahmad, A., & Aboobaider, B. M. (2014, April 23-24). Application of support vector machine for classification of multispectral data. In *7th IGRSM International Remote Sensing & GIS Conference and Exhibition* (p. 012038). Kuala Lumpur, Malaysia.

Chejarla, V. R., Maheshuni, P. K., & Mandla, V. R. (2016). Quantification of LST and CO_2 levels using Landsat-8 thermal bands on urban environment. *Geocarto International*, *31*(8), 913-926.

Cilliers, E. J. (2015). The Importance of Planning for Green Spaces. *Agriculture, Forestry and Fisheries, 4*(4-1), 1-5.

Crommelinck, S., Bennett, R., Gerke, M., Nex, F., Yang, M., & Vosselman, G. (2016). Review of automatic feature extraction from high-resolution optical sensor data for UAV-based cadastral mapping. *Remote Sensing*, *8*, 1-28.

Dewan Bandaraya Kuala Lumpur. (2013). *Draft Kuala Lumpur City Plan 2020*. Kuala Lumpur, Malaysia: Author.

Dewan Bandaraya Kuala Lumpur. (2018). *Kuala Lumpur City Plan 2020.* Kuala Lumpur, Malaysia: Author

De Ridder, K., & Lefebre, F. (2003). BUGS, Benefits of Urban Green Space. *MESO WP final report*.

Elsayed, I. S. M. (2012). Mitigation of the urban heat island of the city of Kuala Lumpur, Malaysia. *Middle-East Journal of Scientific Research*, *11*(11), 1602-1613.

Government of Malaysia Performance Management and Delivery Unit. (2016). *National Transformation Programmed Annual Report 2016*. Putrajaya, Malaysia: Author.

Hsu, C. W., Chang, C. C., & Lin, C. J. (2010). *A practical guide to support vector classification* [Lecture PowerPoint slides]. Retrieved February 19, 2019, from https://www.csie.ntu.edu.tw/~cjlin/talks/freiburg.pdf

Isa, N. A., Wan Mohd, W. M. N., & Salleh, S. A. (2017, October 4). The effect of built-up and green areas on the land surface temperature of the Kuala Lumpur City. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* (Vol. XLII, pp 107-112). Kuala Lumpur, Malaysia.

Jiang, J., & Tian, G. (2010, August 27-29). Analysis of the impact of land use / land cover change on land surface temperature with remote sensing. In *International Society for Environmental Information Sciences 2010 Annual Conference (ISEIS)* (pp 571-575). Beijing, China.

Karuppannan, S., Baharuddin, Z. M., Sivam, A., & Daniels, C. B. (2014). Urban green space and urban biodiversity: Kuala Lumpur, Malaysia. *Journal of Sustainable Development, 7*(1), 1-16.

Kaur, D., & Kaur, Y. (2014). Various image segmentation techniques : a review. *International Journal of Computer Science and Mobile Computing*, *3*(5), 809-814.

Kim, H., Lee, D. K., & Sung, S. (2016). Effect of urban green spaces and flooded area type on flooding probability. *Sustainability (Switzerland), 8*(2), 134-150.

Kwak, Y. (2016). *GIS-Base Suitability Analysis and Planning of Green Infrastructure: A Case of the PPCOD, Capitol Hill* (Master thesis). University of Washington, US. Retrieved December 10, 2018, from http://hdl.handle.net/2123/7128

Li, X., & Shao, G. (2014). Object-based land-cover mapping with high resolution aerial photography at a county scale in midwestern USA. *Remote Sensing*, *6*(11), 11372-11390.

Liedtke, J., & Hu, H. (2018). ArcGIS Pro : Image Segmentation , Classification , and Machine Learning (27). *Esri User Conference*. San Diego, California.

Liu, H., & Shen, Y. (2014). The impact of green space changes on air pollution and microclimates: A case study of the Taipei Metropolitan area. *Sustainability, 6*, 8827-8855.

Mansor, M., & Said, I. (2008). Green infrastructure network as social spaces for well-being of urban residents in Taiping , Malaysia. *Jurnal Alam Bina*, *11*(2), 1-18.

Michel, J., Youssefi, D., & Grizonnet, M. (2015). Stable mean-shift algorithm and its application to the segmentation of arbitrarily large remote sensing images. *International Transactions on Geoscience and Remote Sensing*, *53*(2), 952-964.

Moskal, L. M., Styers, D. M., & Halabisky, M. (2011). Monitoring urban tree cover using object-based image analysis and public domain remotely sensed data. *Remote Sensing*, *3*(10), 2243-2262.

Nakata-Osaki, C. M., Souza, L. C. L., & Rodrigues, D. S. (2018). THIS – Tool for Heat Island Simulation: A GIS extension model to calculate urban heat island intensity based on urban geometry. *Computers, Environment and Urban Systems*, *67*, 157-168.

Nor, A. N. M., Corstanje, R., Harris, J. A., & Brewer, T. (2017). Impact of rapid urban expansion on green space structure. *Ecological Indicators*, *81*, 274-284.

Noorollahi, Y. (2005). *Application of GIS and Remote Sensing in Exploration and Environmental Management of Namafjall Geothermal Area, N-Iceland* (Master thesis). University of Iceland, Iceland.

Rakhshandehroo, M., Mohdyusof, M. J., Tahir, O. M., & Yunos, M. Y. M. (2015). The social benefits of urban open green spaces : a literature review. *Management Research and Practice*, *7*(4), 60-71.

Rudrapal, D., & Subhedar, M. S. (2015). Land cover classification using support vector machine. *International Journal of Engineering Research and Technology (IJERT)*, *4*(09), 584-588.

Santos, T., Tenedório, J. A., & Gonçalves, J. A. (2016). Quantifying the city's green area potential gain using remote sensing data. *Sustainability (Switzerland)*, *8*(12), 1-16.

Shi, D., & Yang, X. (2015). Support vector machines for land cover mapping from remote sensor imagery. In J. Li & X. Yang (Eds.), *Monitoring and Modeling of Global Changes: A Geomatics Perspective* (pp. 265-279). Dordrecht, Netherlands: Springer.

Shubho, M. T. H., Islam, S. R., Ayon, B. D., & Islam, I. (2015). An improved semiautomatic segmentation approach to land cover mapping for identification of land cover change and trend. *International Journal of Environmental Science and Technology*, *12*(8), 2593-2602.

Taati, A., Sarmadian, F., Mousavi, A., Pour, C. T. H., & Shahir, A. H. E. (2015). Land use classification using support vector machine and maximum likelihood algorithms by landsat 5 TM images. *Walailak Journal of Science and Technology*, *12*(8), 681-687.

Tani, A., & Chang, J. (2017). Air quality and health effects of biogenic volatile organic compounds emissions from urban green spaces and the mitigation. *Environmental Pollution*, *230*, 849-861.

Vatseva, R., Kopecka, M., Otahel, J., Rosina, K., Kitev, A., & Genchev, S. (2016, June 13-17). Mapping urban green spaces based on remote sensing data: case studies in Bulgaria and Slovakia. In *Proceedings of 6th International Conference on Cartography and GIS* (pp. 569-578). Albena, Bulgaria.

Wężyk, P., Hawryło, P., Szostak, M., & Pierzchalski, M. (2016). Using geobia and data fusion approach for land use and land cover mapping. *Quaestiones Geographicae*, *35*(1), 93-104.

Wolch, J. R., Byrne, J., & Newell, J. P. (2014). Urban green space, public health, and environmental justice: The challenge of making cities 'just green enough'. *Landscape and Urban Planning*, *125*, 234-244.

Yaakup, A., Zalina, S., Bakar, A., & Bajuri, H. (2005, November 21). GIS based integrated planning assessment for sustainable land use development. In *National Planning Seminar, Selangor, Malaysia* (pp. 1-14). Shah Alam, Malaysia.

Yusof, M. J. M. (2012). Identifying green spaces in Kuala Lumpur using higher resolution satellite imagery. *Alam Cipta*, *5*(2), 93-106.

Pertanika J. Sci. & Technol. 27 (4): 1991 - 2013 (2019)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Computational Study of Mixed Heat Convection in Annular Space between Concentric Rotating Inner and Wavy Surface Outer Cylinders

Qusay Rasheed Al-Amir*, Farooq Hassan Ali Alinnawi and Qusay Adnan

Department of Mechanical Engineering, Faculty of Engineering, University of Babylon, Babylon Province, Al Hillah, Babylon City, Iraq

ABSTRACT

Mixed heat convection inside annular spaces occurs in many engineering technology applications. This study aims to determine the effect of the sinusoidal surface parameters of an outer cylinder, which are represented by variations in the undulation number and amplitude of the wavy surface, on flow structure and thermal fields for different values of the Reynolds number (Re; from 0 to 600) and Rayleigh number (Ra; from 10^3 to 10^6). A horizontal annular space bounded by two concentric cylinders contained air with a Prandtl number that equaled 0.7. The sinusoidal surface of the fixed circular outer cylinder was maintained at a constant cold temperature (T_c), whereas the surface of the circular inner cylinder was set at a constant hot temperature (T_h) and rotated in counter-clockwise direction at constant angular velocity. Calculations were performed under steady-state conditions. A computational procedure based on the finite volume technique was implemented using the software ANSYS Fluent (version 16.1). Results indicate that the heat transfer from the inner cylinder increased with a rise in the surface amplitudes and undulation numbers with a fixed Re. The average Nusselt number increased with an increase in Ra and reduced when the undulation number increased from one to two. In summary, the heat transfer of the cylinder with the sinusoidal outer surface is better by 7.3% than that of the conventional cylinder.

Keywords: Annular space, concentric cylinders, mixed convection, rotating cylinder, surface amplitude, undulation number

ARTICLE INFO

Article history:
Received: 11 March 2018
Accepted: 23 April 2019
Published: 21 October 2019

E-mail addresses:
qusay1972@gmail.com (Qusay Rasheed Al-Amir)
farooq_hassan77@yahoo.com (Farooq Hassan Ali Alinnawi)
qam_me782000@yahoo.com (Qusay Adnan)
* Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

INTRODUCTION

Mixed convection heat transfer within annular spaces has been a topic of interest for decades. The mixed convection problem in finite spaces is present in several industrial and environmental applications, including the ventilation of buildings, waste transport and storage of nuclear and chemical reactors, design of electronic equipment (Mozayyeni & Rahimi, 2011), cooling systems (Yoo, 1998), different types of solar collectors and thermal storage systems (Bande & Mariah, 2014) and thermal management in aviation (Xu et al., 2009). Moreover, mixed convection is becoming increasingly important in engineering applications, especially because enclosures with internal bodies are more complicated than classical enclosures. Mixed convection in different enclosure geometries, such as square, rectangle and cylinder, has been investigated in various studies (Shih et al., 2009; Chamkha et al., 2011; Hussain & Abd-Amer, 2011; Wang et al., 2015; Rajamohan et al., 2015). These studies experimentally, analytically and numerically investigated mixed convection in annular spaces under different boundary conditions.

Sheikholeslami et al. (2012) developed a FEM model of the natural convection heat transfer in an annular space bounded by a circular outer enclosure and a sinusoidal circular inner cylinder. Their study was conducted at different dimensionless groups, such as the Hartmann number (Ha), Rayleigh number (Ra), volume fraction of nanoparticles and undulation number of the inner cylinder. The authors revealed that the $(\overline{Nu})$ variation was a function that increased with the volume fraction of nanoparticles, the number of undulations and Ra but decreased with Ha. Sheikholeslami et al. (2013) studied the effect of the amplitude and the undulation number of a hot sinusoidal inner cylinder on the natural convection problem with the assumption that the enclosure (circular outer) was filled with air and maintained at a cold temperature. Their works covered a wide range of parameters (Ra = 10^3, 10^4, 10^5 and 10^6), amplitudes (L = 0.1, 0.3 and 0.5) and numbers of undulations of the inner cylinder (N = 2, 3, 5 and 6). These authors found that the flow and temperature patterns and cells (number, size and formation) strongly depended on the Ra, undulation number and amplitude of the enclosure. Magneto hydrodynamic flow in a nanofluid-filled inclined enclosure was numerically investigated by Sheikholeslami et al. (2014), who concluded that the velocity field retarded and the convection and Nusselt number (Nu) decreased in the existence of the magnetic field. Shekar et al. (1984) numerically studied the convective heat transfer and flow characteristics of a fluid filling the space between two horizontal concentric cylinders. The results were obtained under the effects of an externally imposed temperature gradient across the annulus and uniform internal heat generation. The authors concluded that the ratio of the characteristic temperature had an important role in the flow behaviour, in which the flow fields comprised one or two vortices in each half cavity. Xu et al. (2010) conducted laminar natural convective heat transfer around a horizontal cylinder inside a concentric triangular enclosure. The researchers concluded that the predicted overall Nu of an inclination angle was independent of the inclination angle. Laminar natural heat transfer around a coaxial triangular inner cylinder to its concentric cylindrical enclosure was performed by Yu et al. (2010). The results revealed that the temperature distribution became nearly independent of the Prandtl number. Yuan et al. (2015) analysed the free convection in the horizontal concentric annuli problem with different inner cylinders. They concluded that the surface radiation and presence of corners and large top space played an important role in enhancing the heat transfer rate.

The mixed convection in the entry region of a vertical annulus with a rotating inner cylinder was presented by El-Shaarawi and Sarhan (1982). Their results indicated that the hydrodynamic development length and heat transfer parameters were affected by the rotating inner cylinder. Yoo (1998) investigated numerically the mixed convection heat transfer characteristics of air within two horizontal concentric cylinders with different uniform temperatures. Their investigation revealed the flow regimes on the Ra–Re plane, and they presented and explained the characteristics of the velocity patterns and the heat transfer. Mohammed (2007) used a numerical finite difference method to obtain the solution to the mixed convection in the entry region of a vertical annulus containing an inner cylinder. The results showed that the thickness of the thermal boundary layer gradually increased as the flow moved from the annulus inlet towards the annulus exit. Alam et al. (2016) studied mixed convection inside a differentially heated square enclosure containing a rotating heat conducting cylinder. The results indicated that the flow field, temperature distribution and rate of heat transfer were dependent on cylinder size and rotating speeds. Alsabery et al. (2018) studied the effect of rotating solid cylinder on entropy generation and convective heat transfer in a wavy porous cavity heated from below. They showed that the heat transfer rate from the wall to the fluid increased with the Darcy number and, in particular, increased significantly for $Da > 10^{-3}$.

A few experimental studies on the natural convective heat transfer across an annulus with a hexagonal inner cylinder and a concentric circular outer cylinder have been conducted, such as that by Boyd (1984). A correlation for the mean Nu at the surface of the inner cylinder was presented for a specific range of Ra and expressed as $Nu = 0.794\ Ra^{0.25}$. The results revealed that the presence of hexagonal inner element corners enhanced the average Nu compared with the heat transfer in a circular inner cylinder. Lee (1992) conducted an experimental and numerical study on the effect of the convective fluid motion of air enclosed between the annuli of inner and eccentric horizontal cylinders, which were assumed to be heated and rotating. The influence of the cylinder through its rotating speed on the resulting convection problem was investigated. The result showed that the higher the rotational Re, the more the flow tended to become uniform. Furthermore, the inner rotational cylinder remarkably affected the mean Nu. Another experimental work was presented by Rajamohan et al. (2015) to explore the effect of various inclination angles on the mixed convection heat transfer of a thermally developing flow in a side-heated square duct. The authors found that an increase in the inclination angle improved the convection rate and hence remarkably enhanced the heat transfer.

As previously mentioned in the literature, mixed convection in annular spaces bounded with one corrugated surface has not been explored. The present work aims to investigate the mixed convection heat transfer in the annular space between concentric cylinders with corrugated outer and smooth inner surfaces. The effect of various parameters, such as undulation number, amplitude of the wavy surface, Re and Ra, is considered. These parameters are investigated in a wide range to examine the isotherms and streamline contours of the annular space.

PHYSICAL MODEL AND COORDINATE SYSTEM

Figure 1 shows a schematic of the two-dimensional concentric cylinders; the outer cylinder is a sinusoidal, whereas the inner one is circular. In this study, the air is utilised as the working fluid, and the ratio between the inner diameter cylinder and the subdivision of the mean radius of the outer and inner radius cylinders ($R_i/(R_{om}–R_i)$) is equal to 2. Therefore, the characteristic length is equal to the space between the two concentric cylinders (i.e. $R_{om}–R_i$). The sinusoidal outer and circular inner cylinders are assumed to have the same centre in the origin of the Cartesian coordinate. The functions are applied to obtain the sinusoidal wavy wall on the surface of the circular cylinder.

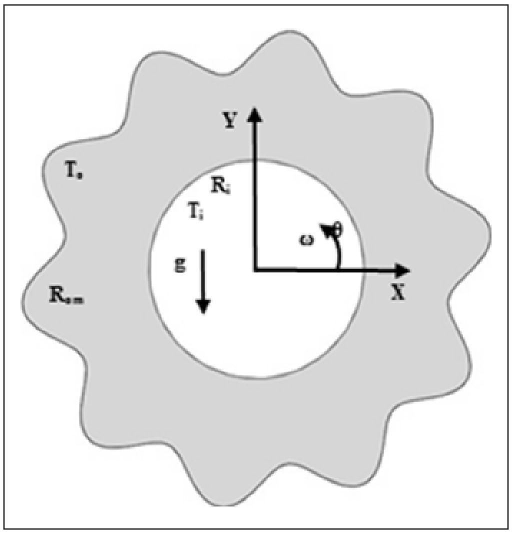


Figure 1. Physical model of the present study

$$X = [R + L sin(n * \theta * \pi/180)/n] * \left(cos\left(\theta * \frac{\pi}{180}\right)\right) \quad [1]$$

$$Y = [R + L\sin(n * \theta * \pi/180)/n] * (\sin\left(\theta * \frac{\pi}{180}\right)) \quad [2]$$

where X and Y are the Cartesian coordinates, R is the relaxation function, L is the amplitude of the wavy surface, n is the undulation number and θ is the angle of the circular cylinder. The cylinder length is assumed infinite where the length to diameter ratio is very long. The outer and inner cylinders are kept at cold and hot temperatures, respectively, and the difference between them is fixed at 50°. The difference in the temperature that produces the buoyancy effect leads to natural convection, whereas the rotation of the inner cylinder with rotating velocity (ω) leads to forced convection. Hence, the combination of these conditions produces mixed convection effects.

GOVERNING EQUATIONS AND BOUNDARY CONDITIONS

In the present study, the software ANSYS Fluent version 16.1 (ANSYS, 2014) was used to solve the governing equations. The control volume technique was used to transfer the governing equation to algebraic equations. The numerical solutions were based on the control volume approach with the integration of the governing equation in each control volume and then discretization of the equations that were transfered from any quantity to the control volume. The governing continuity equation and momentum and energy equations for laminar, incompressible, two-dimensional and steady state can be written with following dimensionless variables.

$$X = \frac{x}{R_c}, Y = \frac{y}{R_c}, U = \frac{u}{u_0}, V = \frac{v}{v_0}, P = \frac{p}{\rho u_0^2}, \emptyset = \frac{T - T_c}{\Delta T}, \Psi = \frac{\psi}{u_0 R_c}, Pr = \frac{\nu}{\alpha},$$

$$Re = \frac{u_0 R_c}{\nu}$$

Continuity Equation:

$$\frac{\partial U}{\partial x} + \frac{\partial V}{\partial y} = 0 \qquad [3]$$

Momentum Equations:

$$U\frac{\partial U}{\partial X} + V\frac{\partial V}{\partial Y} = -\frac{\partial P}{\partial X} + \frac{1}{Re}\left(\frac{\partial^2 U}{\partial X^2} + \frac{\partial^2 U}{\partial Y^2}\right) \qquad [4]$$

$$U\frac{\partial V}{\partial X} + V\frac{\partial V}{\partial Y} = -\frac{\partial P}{\partial Y} + \frac{1}{Re}\left(\frac{\partial^2 V}{\partial X^2} + \frac{\partial^2 V}{\partial Y^2}\right) + \frac{Ra}{PrRe^2}\emptyset \qquad [5]$$

Energy Equation:

$$U\frac{\partial \emptyset}{\partial X} + V\frac{\partial \emptyset}{\partial Y} = \frac{1}{RePr}\left(\frac{\partial^2 \emptyset}{\partial X^2} + \frac{\partial^2 \emptyset}{\partial Y^2}\right) \qquad [6]$$

where ρ is the fluid density, u and v are the Cartesian velocity components, u_o and v_o are reference Cartesian velocity components, Ψ is dimensionless stream function, P is the dimensionless pressure, g is the gravitational acceleration, μ is dynamic viscosity, T is the temperature and α is the thermal diffusivity. In the present work, air is used as the working fluid and has a Prandtl number of 0.71. The reference temperature is set to 300 K$(T_r = \frac{T_o + T_i}{2})$. Table 1 lists the properties of air of this reference temperature. In addition, Ra is defined as

$$Ra = \frac{g\beta\Delta T R_c^3}{\nu^2} \qquad [7]$$

where T_h and T_c are the hot and cold temperatures, respectively; l is the length of the annulus; υ is the kinematic viscosity and β is the thermal expansion ($\beta = \frac{1}{T_r}$), which changes with Ra. The boundary conditions are assumed as follows.

At the outer cylinder surface, To = 275 K and u = v = 0.

At the inner cylinder surface, T_i = 325 K, u = ω r sinθ and v = ω r sinθ.

The local Nusselt number (Nu_s) can be calculated to estimate the heat transfer rate along the hot inner cylinder surface. It is a non-dimensional term that is defined as follows:

$$Nu_s = \frac{h.ds}{k_r} \quad [8]$$

Where h is convective heat transfer coefficient (W m^{-2} K^{-1}), d_s is inner circular cylinder perimeter (m), and k_r is thermal conductivity W m^{-1} K^{-1}. In the analysis, the average Nusselt number ($\overline{Nu}$) is defined as

$$\overline{Nu} = \frac{1}{2\pi}\int_0^{2\pi} Nu.ds \quad [9]$$

We adopted a mapped mesh, with a number of divisions equal to 140 for inner, outer cylinder (Edge sizing) and the space between them (face meshing) which produced 19740 nodes and 19600 elements to describe the heat transfer and fluid flow behaviour between the two concentric cylinders, as shown in Figure 2. A grid check with specific conditions, namely, number of corrugations = 15, amplitude (L) = 0.5, Ra = 10^5, Re = 118.678, was tested to obtain a suitable number of elements with high accuracy (i.e. continuity = 10^{-6}, V_x = 10^{-6}, V_y = 10^{-6} and energy = 10^{-8}) (Figure 3). The convergence criteria discretization of the CFD solution was supposed for the relative error for any dependent variables satisfies the discretisation convergence criteria is $\left|\frac{\zeta_{i+1} - \zeta_i}{\zeta_{i+1}}\right| \leq \Gamma$. Where ($i$) signifies the number of iterations and (Γ) is the discretization convergence criteria. Many grid independent checks were examined to evaluate the grid sensitivity. Table 2 shows the grid independence check and percentage error. The error is plotted as a function of iteration number are shown in Figure 4. The value of Re=118.678 is corresponding to the Richardson number Ri=10 which represents the middle value between pure natural convection (Ri= ∞) and pure forced convection (Ri=0). The reference values for velocity (angular velocity ω=2.35793 rad/s, density (ρ=1.1614 kg/m^3), dynamic viscosity (μ=1.846*10^{-5} m^2/s). The values of Rayleigh number are Ra=10^3, 10^4 and 10^5. The reference quantities for Rayleigh number are (α=2.25*10^{-5} m^2/s), (υ=1.589*10^{-5} m^2/s), the temperature difference ΔT=50, characteristic length (r=0.02 m), gravitational acceleration (g=9.81 m^2/s) and β is varied corresponding to Rayleigh number. No remarkable change occurs in the average Nu when the element number reaches 19,600.

Table 2
Error discretization of grid independence for average Nusselt number along the hot inner cylinder for N=15, L=0.5, Ra=10⁵, Re=118.678

Number of Divisions $\frac{\boldsymbol{Outer\ cylinde * Inner\ cylinde}}{\boldsymbol{Face\ meshing}}$	Number of elements	Mean Nusselt number at the hot wall	Error (%)
$\frac{\mathbf{40 * 40}}{\mathbf{40}}$	1600	1.177	-
$\frac{\mathbf{60 * 60}}{\mathbf{60}}$	3600	1.186	0.7862
$\frac{\mathbf{80 * 80}}{\mathbf{80}}$	6400	1.187	0.0378
$\frac{\mathbf{120 * 120}}{\mathbf{120}}$	14400	1.191	0.3093
$\frac{\mathbf{140 * 140}}{\mathbf{140}}$	19600	1.191	0.0014

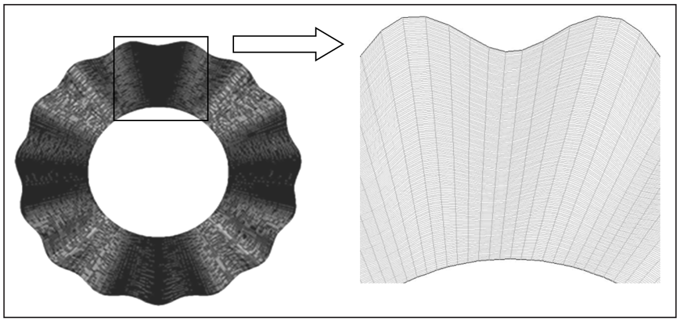


Figure 2. Mapped mesh of the present study

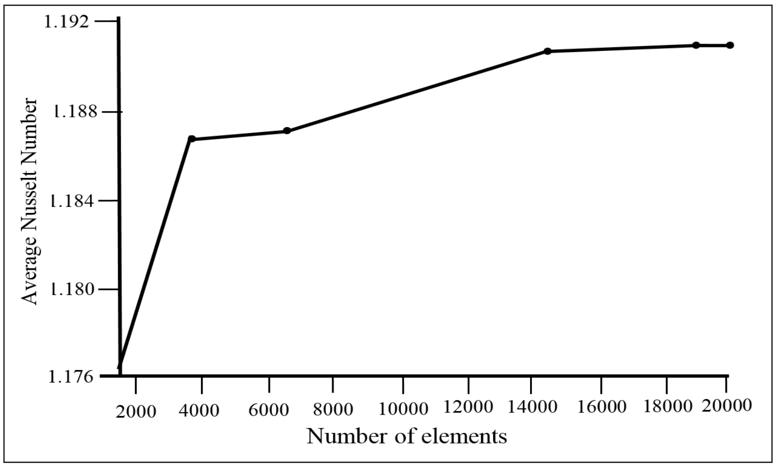


Figure 3. Verification of mesh generation

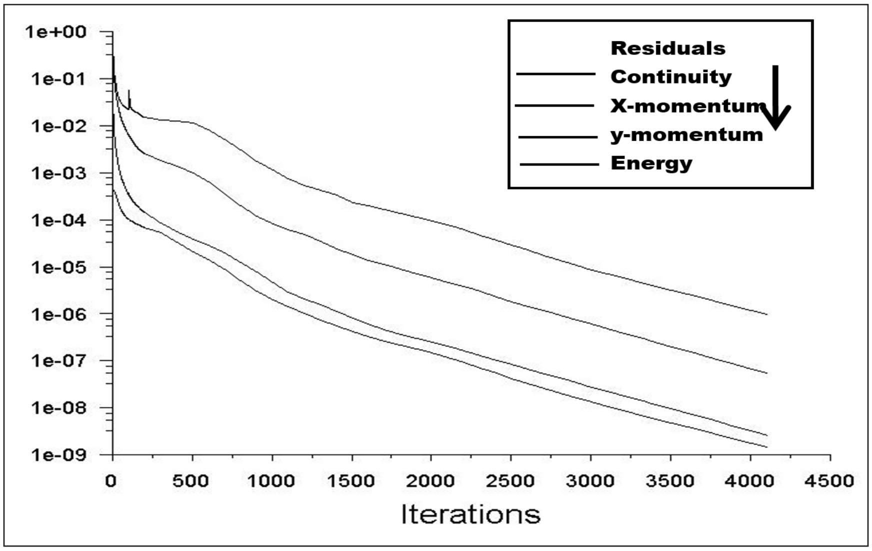


Figure 4. Error as a function of iteration number

VALIDATION

To validate the solution procedure, the numerical results obtained for both the streamlines and isotherms contours in an air-filled annular space between a heated inner concentric rotating cylinder and wavy surface outer cylinder were compared with results published by Yoo (1998). The comparison was done using the following parameters: Ra = 10^5, Pr = 0.71, at l/D=2.5, Re=100, 300, 400, and 600, respectively. As can be seen in the Figure 5, close agreements were observed between the present results and those presented in Yoo (1998).

RESULTS AND DISCUSSION

The numerical investigations in this study were conducted under the following parameter ranges: amplitudes of the wavy surface (L = 0, 1 and 2), undulation numbers (N = 5, 10 and 15), Re (Re = 0, 10, 100, 200, 400 and 600) and Ra (Ra = 10^3, 10^4 and 10^5). Numerous results could be drawn from the output of ANSYS Fluent version 16.1 (ANSYS, 2014), such as pressure, flow trace, velocity vectors and streamline and isotherm contours. However, given the space limitation, only the results of the streamline and isotherm contours are shown in Figures 6(a–b) to 8(a–b). In addition, the heat transfer rate along the hot surface of the circular cylinder is expressed by local and average Nu, as shown in Figures 9(a–c) and 10.

Flow Structure and Thermal Fields

The flow structure and thermal fields were visualised using streamlines and isotherms, respectively. The streamline and isotherm patterns in Figures 6(a) and 6(b) demonstrated the effects of the amplitudes of the wavy surface, undulation numbers and Re on the flow

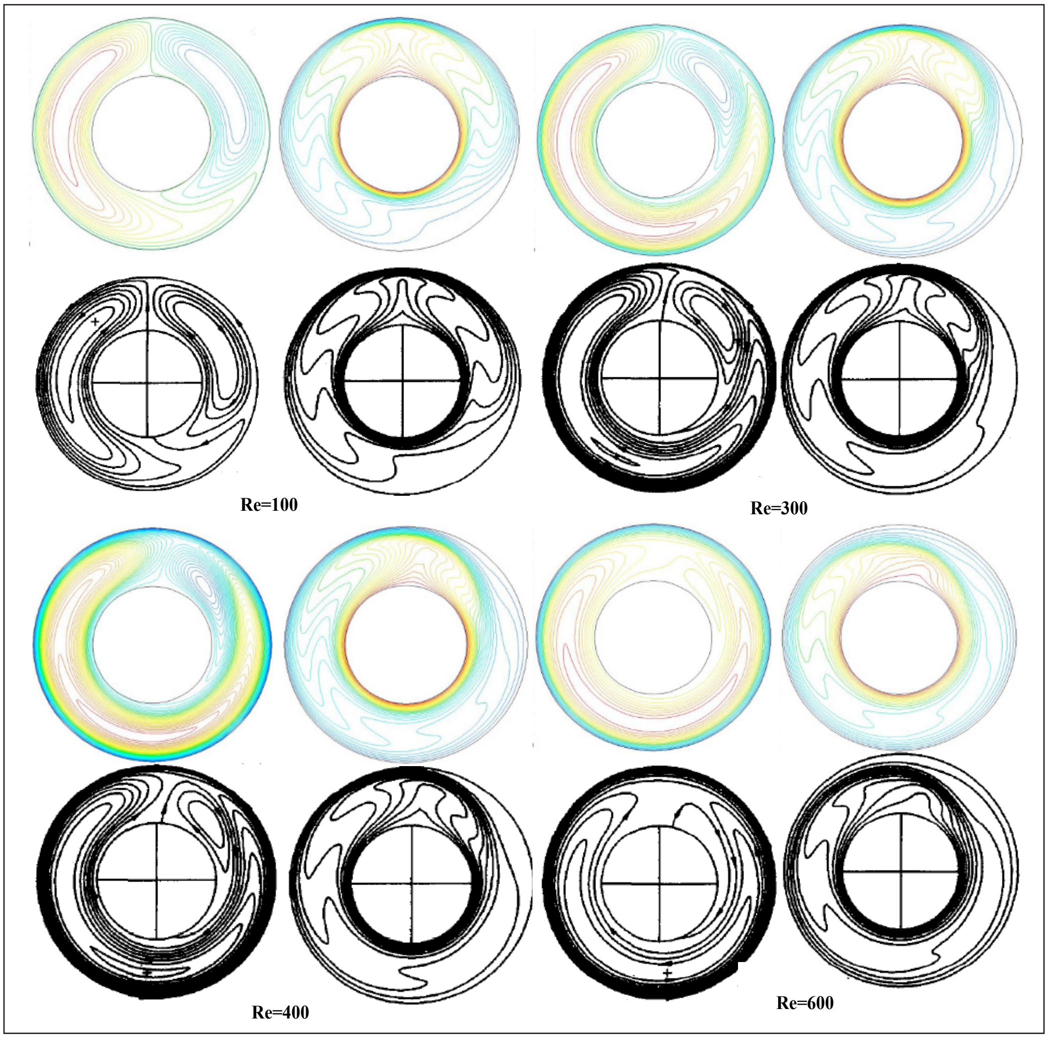


Figure 5. Comparison of streamline, isothermal line and heat transfer for present study and results of Yoo (1998) for l/D=2.5, Pr=0.71, Ra=10^5, Re=100, 300, 400, and 600, respectively

structure and thermal fields for a fixed Ra of 10^3. When the flow was at a low Ra of 10^3 and Re was equal to 0, the fluid motion inside the annulus was weak. The conduction or diffusion mode dominated the heat transfer between the hot and cold cylinders. In each selected case, the streamlines showed two rotating symmetric eddies; i.e. one on the right hand and the other on the left hand of the annular gap. These eddies were compressed and changed in the inner vortices inside each eddy possibly due to the wavy surface of the outer cylinder. Meanwhile, the isotherms were crowded around the hot inner cylinder and form parallel rings around the cylinder surface. Moreover, the contours of the streamlines and isotherms were somewhat systematic for those corresponding to Re = 10. As the value of Re increased to 100, the rotational speed of the inner cylinder led to a large change in the

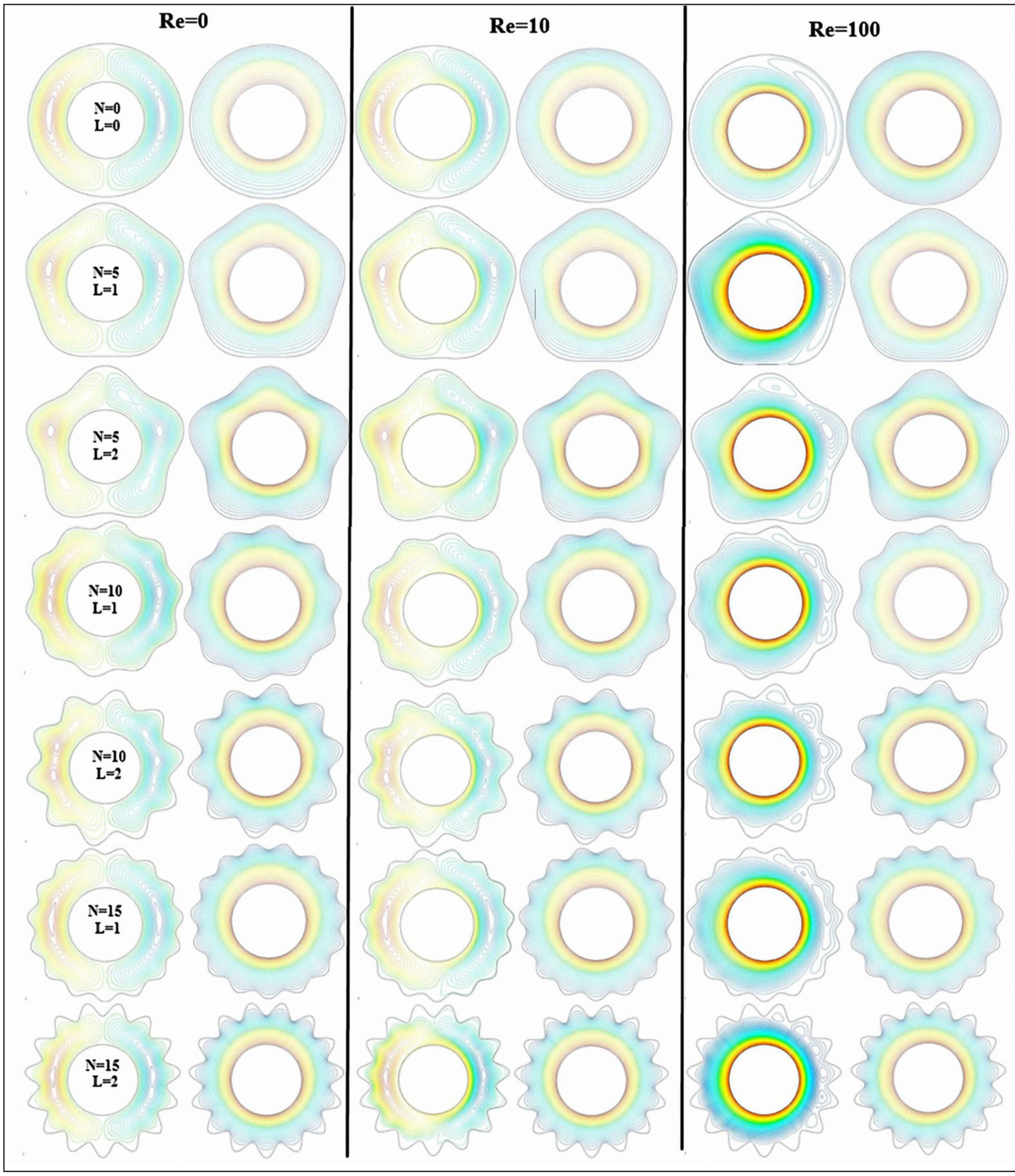


Figure 6a. Streamlines on left, Isothermal lines on right for Different Numbers of sinusoidal (N), amplitude (L) and Reynolds Number (Re=0,10,100) at Ra=10^3

flow field; hence, the streamlines were remarkably dense around the inner cylinder, and the two rotating eddies were progressively compressed to a single longitudinal eddy in the vicinity of the right side of the annulus gap case (N = 0, L = 0). In the other cases, the eddy was pushed to the wavy surface of the outer cylinder due to the centrifugal force and spread to the number of cells (i.e. three vortex cells for case [N = 5, L = 2]; five vortex cells for case [N = 10, L = 2]; and six vortex cells for case [N = 15, L = 2]).

The isotherm contours were similar to those of the previous values of Re. The effect of the rotation appeared to be strong such that the small eddies gradually disappeared with an increase in the Re to 200. This condition tends to enhance the influence of the flow induced by forced convection. When Re was increased from 400 to 600, the forced convection dominates the heat transfer and fluid flow. Consequently, the streamlines became concentric rings around the heated cylinder for each case. Conversely, the streamlines disappeared at the outer surface of the cylinder, where the effects of the cylinder rotation were weak. Meanwhile, the isotherms became more evenly distributed between the two cylinders as Re increased.

Figures 7(a) and 7(b) present the effects of the amplitudes of the wavy surface, undulation numbers and Re on the streamlines and isotherms at Ra = 10^4. This Ra value caused an increase in the buoyancy forces inside the annulus gap. For small values of Re (i.e. 0–10), the streamline contours exhibited two rotating eddies, i.e. developing in the right and left halves of the annulus gap. Furthermore, the rising hot fluid separated at the top of the inner cylinder to form a thermal plume (i.e. near $\theta = 90°$). This phenomenon indicates that the inner thermal boundary layer begins to separate from the inner cylinder before reaching the top of the annulus. However, at (N = 10, L = 2), two small plumes appeared (i.e. nearly $\theta \approx 45°$ and $\theta \approx 125°$) in place of the usual single big plume. The two plumes yielded two secondary vortices over the upper portion of the annulus because of the increasing values of the wave amplitude. These vortices increased the heat transfer rates between the two cylinders. When Re further increased to 100, the effects of the rotation and the buoyancy appeared. The interaction between the two effects is called mixed convection. In the right part of the annulus, the forced and natural convection flows were combined because they were in the same anti-clockwise direction, but they were opposite in the left part. Therefore, the right cells seemed stronger than the left, thereby leading to the pushing of the cells in the anti-clockwise direction. Consequently, the pair of counter-rotating vortices, as presented in case (N = 10, L = 2), merged with the two main eddies. Then, the isotherms were strongly influenced by the rotation of the cylinder; thus, the thermal plume at the top of the inner cylinder was tilted in the direction of the cylinder rotation (i.e. nearly $\theta = 120°$ for all cases). The patterns of the streamlines and isotherms when Re = 200 were predicted to be the same as those when Re = 100. With a further increase in Re from 400 to 600, the strength of the forced flow increasingly elevates. The thermal plume is progressively tilted at the top the inner cylinder towards the flow direction.

The streamline and isotherm contours are presented in Figures 8(a) and 8(b); these contours are for the same conditions as in Figures 7(a) and 7(b) but at Ra = 10^5. When Ra rose to 10^5, the circulation effect of the eddies increased and the buoyancy force became the dominant mechanism that drove the convection of the fluid. The effect of the forced flow vanished when Re = 0. The fluid motion remained in the region of the natural

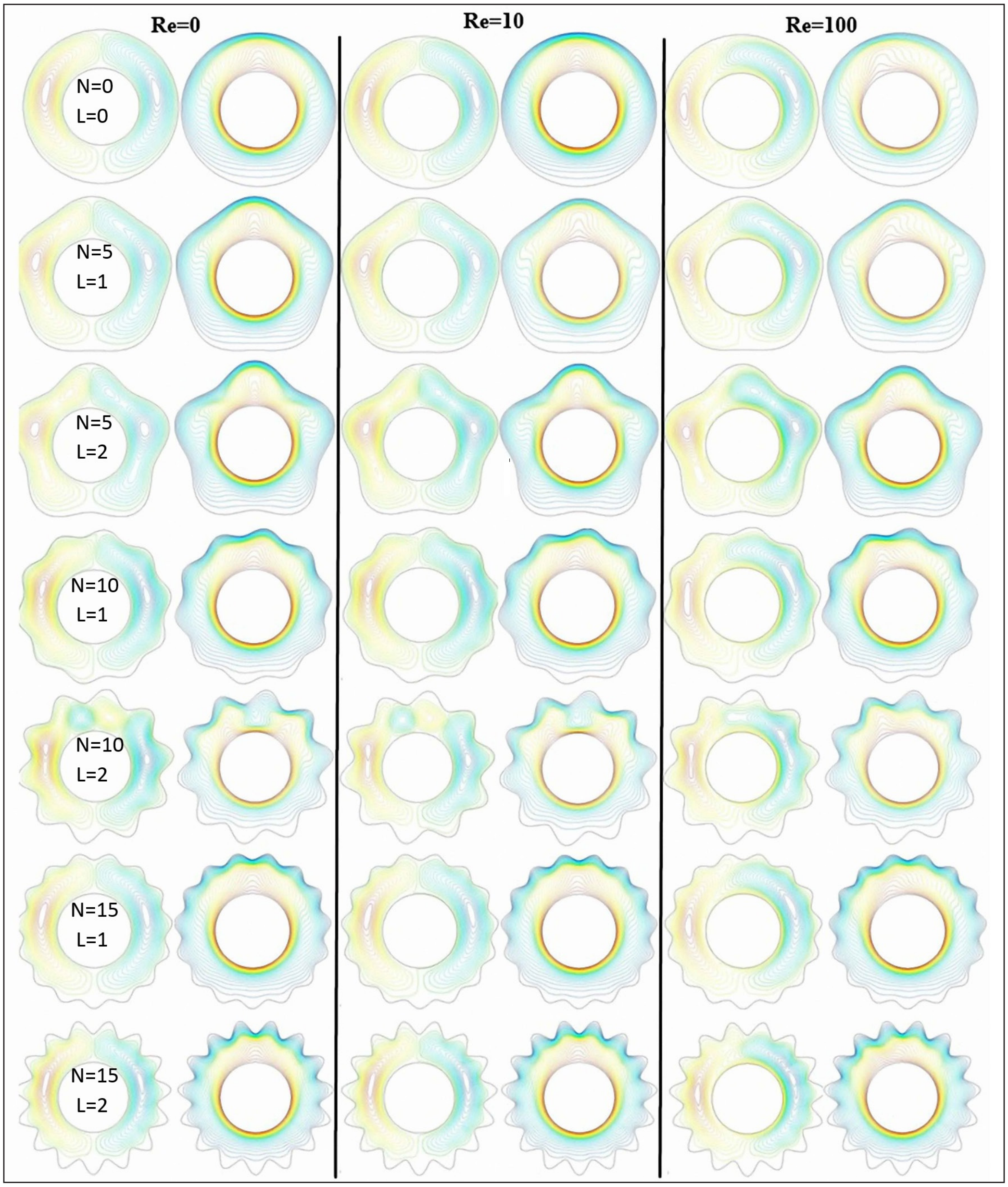


Figure 7a. Streamlines on left, Isothermal lines on right for Different Numbers of sinusoidal (N), amplitude (L) and Reynolds Number (Re) at Ra=10^4

convection, thus causing the formation of the two eddies in each flow pattern. However, the two eddies further rose upwards from the bottom half to the upper half inside the annulus gap due to the buoyancy effect. The flow was generated primarily due to the temperature gradient. As the temperature gradient developed with the increasing Ra, the shape of the isotherms moves upwards to form a thermal plume at the upper part of the annulus (i.e.

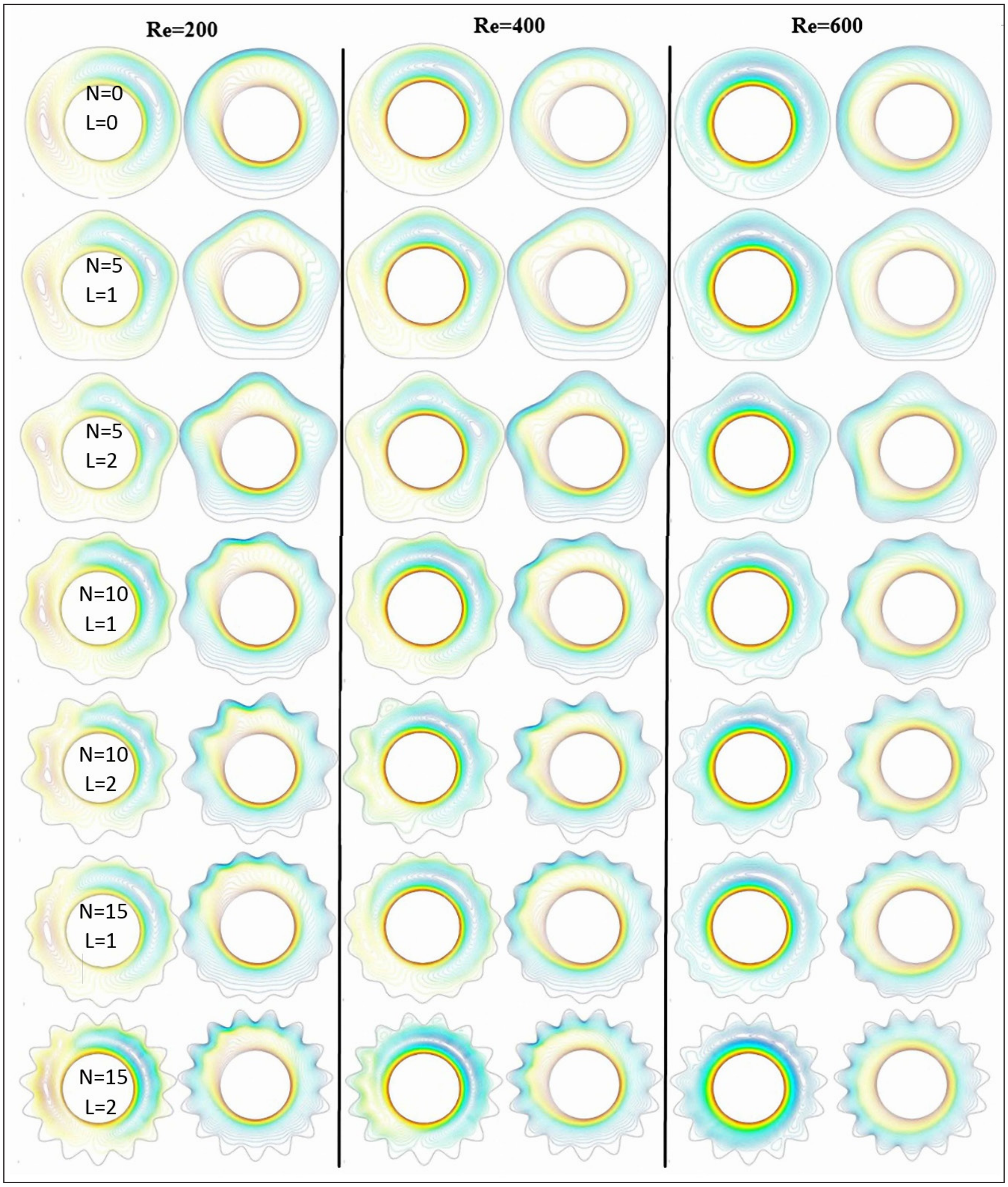


Figure 7b. Streamlines on left, Isothermal lines on right for Different Numbers of sinusoidal (N), amplitude (L) and Reynolds Number (Re=200,400,600) at Ra=10^4

nearly θ ≈ 90°). However, at (N = 15, L = 2), two thermal plumes developed around the angles θ ≈ 45° and θ ≈ 120° over the inner cylinder due to flow separation. In this case, two secondary vortices formed in the presence of the two eddies developing at the upper part of inner cylinder, thereby resulting in a large change in the streamline contour. These vortices produced two thermal plumes that rose to impinge the upper part of the outer

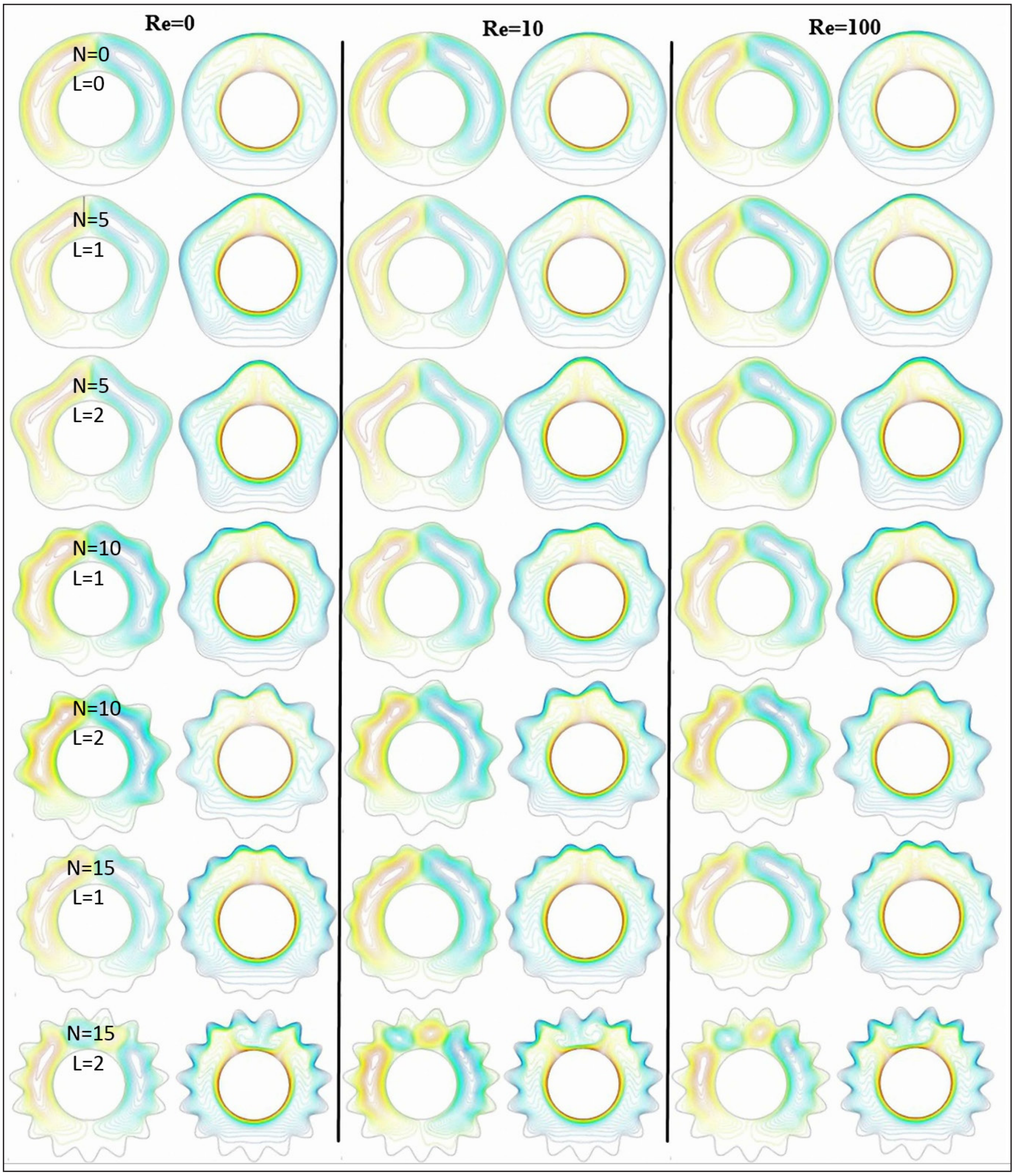


Figure 8a. Streamlines on left, Isothermal lines on right for Different Numbers of sinusoidal (N), amplitude (L) and Reynolds Number (Re=0,10, 100) at Ra=10^5

cylinder. Consequently, the impingement of the plume on the top of the outer cylinder led to the development of a thin thermal boundary layer, thus enabling high heat transfer due to natural convection. A type of symmetry was evident in the patterns of the streamlines and isotherms; this symmetry was similar to that when Re = 10. As Re increased to 100, the effect of the forced flow appeared. The eddy adjacent to the right-hand side of the inner cylinder started to move upwards. Conversely, the eddy on the left-hand side of the

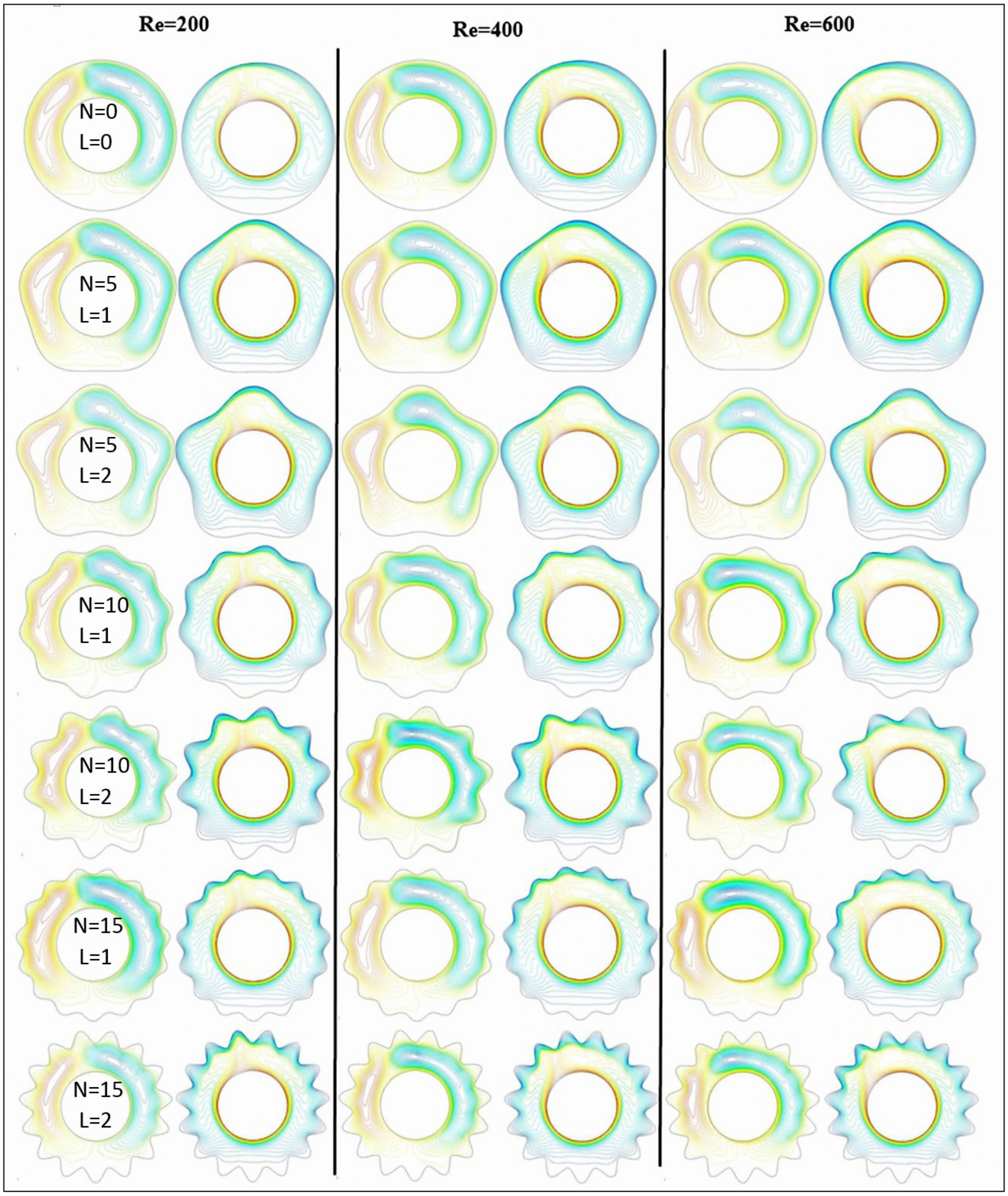


Figure 8b. Streamlines on left, Isothermal lines on right for Different Numbers of sinusoidal (N), amplitude (L) and Reynolds Number (Re=200,400,600) at Ra=10^5

inner cylinder was pulled downwards due to the viscous inner cylinder rotation. Thus, the isotherm patterns depicted that the thermal plume was slanted in the same direction as the rotating cylinder (i.e. nearly $\theta \approx 100°$ for all cases). As Re was increased to 200, the contours of the streamlines and isotherms were shifted continuously towards the rotation of the inner cylinder (i.e. nearly $\theta \approx 115°$ for all cases). In case (N = 15, L = 2), the two small vortices at the top of the inner cylinder disappeared when Re increased to 200. When Re increased from 400 to 600, the shapes of the streamline and isotherm contours in all cases became slightly similar to those at Re = 200, but the thermal plume was slanted constantly in the same direction as the rotating cylinder with increasing Re (i.e. nearly $\theta \approx 135°$ for Re = 400 and $\theta \approx 160°$ for Re = 600).

Heat Transfer and Nu

Table 3 illustrates good agreement between the model result and reported result of Yoo (1998).

Table 3
Comparison of the average Nu between the present and Yoo (1998) results (Re=0, Pr=0.71, l/D=2.5*)*

Ra	Average Nu		
	Present study	Yoo(1998)	Deviation(%)
10^3	0.38189	0.392	1.01
10^4	0.67331	0.737	6.39
10^5	1.20832	1.2601	5.178

Figures 9(a–c) plot the distributions of local Nu values along the hot surface of the inner cylinder for different amplitudes of the wavy surface, undulation numbers and Re at Ra= 10^3, 10^4 and 10^5. When the cylinder was stationary (i.e. Re = 0), the maximum value of the local Nu occurred around the bottom of the cylinder, existing in the vicinity of $\theta \approx 270°$ for Ra = 10^3 and 10^4. Meanwhile, it occurred at the two places between 210° and 315° at Ra = 10^5. In addition, the minimum value of the local Nu was obtained at the top surface of the inner cylinder (i.e. $\theta \approx 90°$). However, two special cases occurred (i.e. N = 10, L = 2 at Ra = 10^4 and N = 15, L = 2 at Ra = 10^5) when a secondary flow was developed at the upper part of the inner cylinder such that the sudden increase in the local heat transfer manifested near the top of the inner cylinder (i.e. $\theta \approx 90°$). In the same figure, the difference between the Nu_{max} and Nu_{min} values increased with the increasing Ra. When the inner cylinder was rotated at Re = 10, the maximum and minimum local Nu values were unaffected due to the small velocity of the inner cylinder. When Re increased to 100, the maximum and minimum values of the local Nu migrated along the rotational direction of the inner cylinder. The variation in the Nu curves became increasingly divergent at Ra = 10^3 when Re increased to 200 and was remarkably small when Ra equaled 10^4 and 10^5. The sinusoidal distributions of the Nu curves became evident at Ra = 10^3 and were suppressed by an increase in the values of Re from 400 to 600**.** Therefore, the lines of the local Nu curves in this Re range strongly depended on the wavy surface of the outer cylinder and thus appearing to be wavy also. This behaviour disappeared at high Ra values (10^4 and 10^5) due to the boundary layer separation.

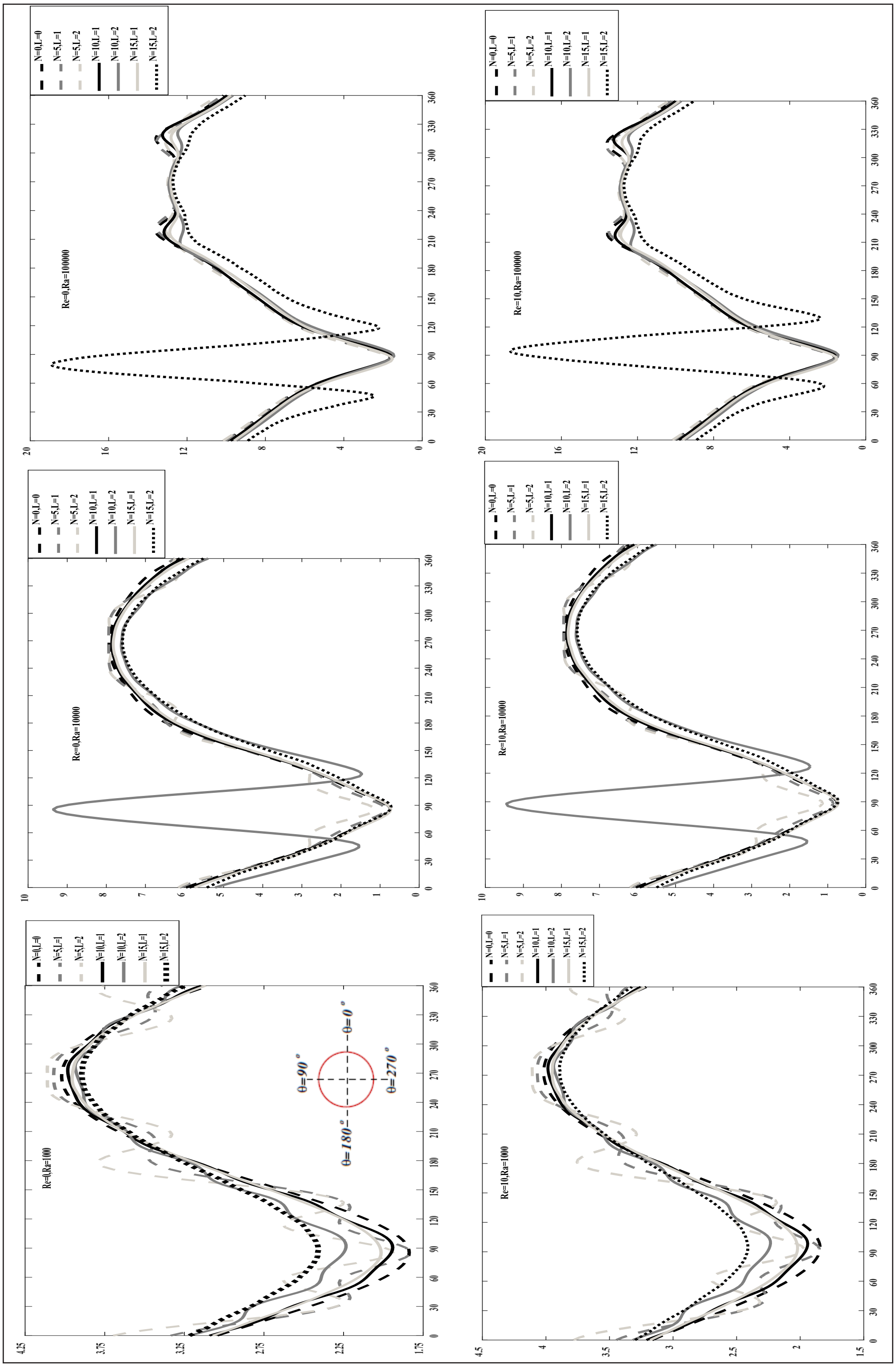


Figure 9a. Local Nusselt Number (y-axis) versus the inner cylinder surface (x-axis) at Re=0, and10 for different Ra=10^3, 10^4 and 10^5, respectively

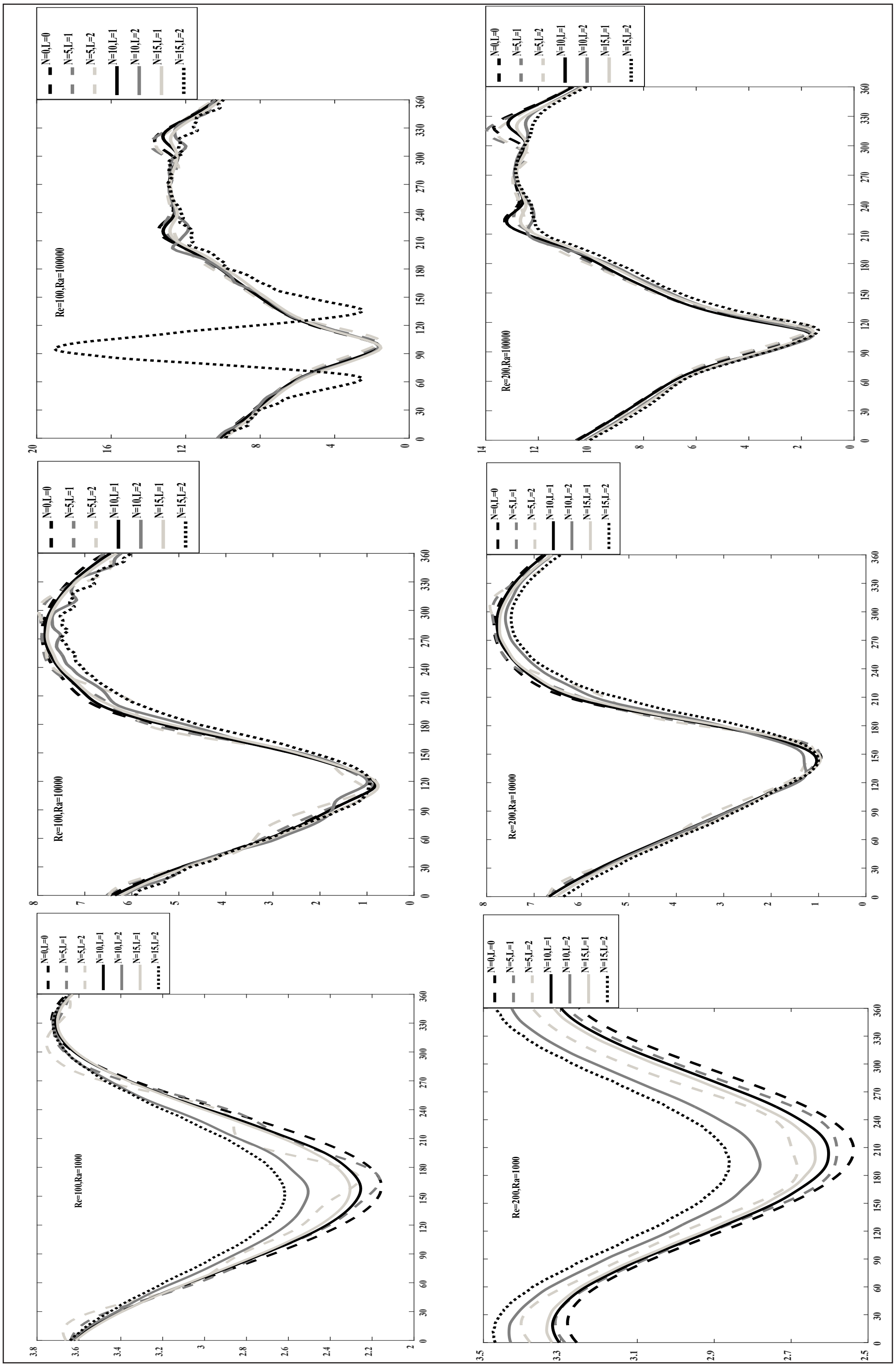


Figure 9b. Local Nusselt Number (y-axis) versus the inner cylinder surface (x-axis) at Re=100, and 200 for different Ra=10^3, 10^4 and 10^5, respectively

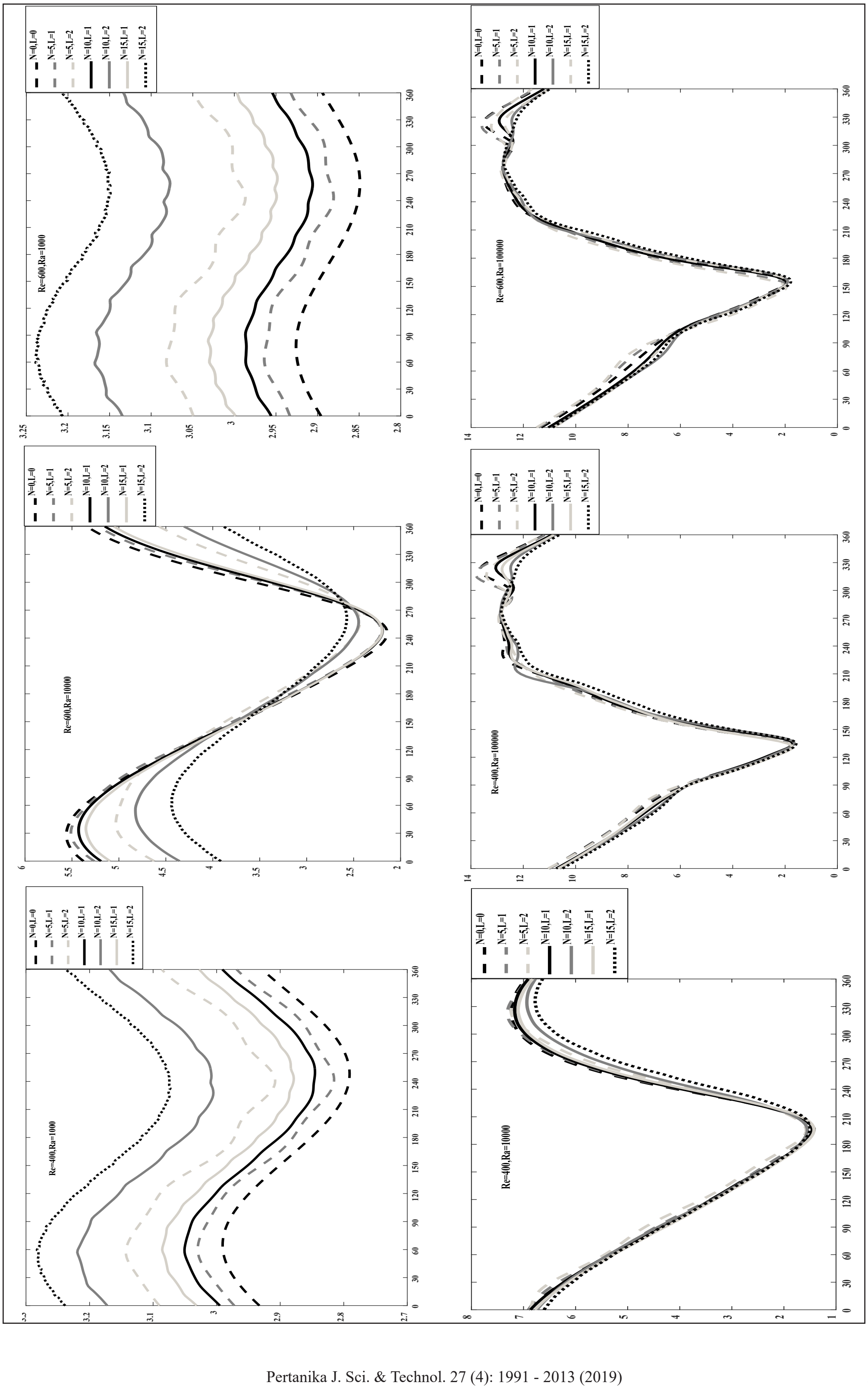


Figure 9c. Local Nusselt Number (y-axis) versus the inner cylinder surface (x-axis) at Re=400, and 600 for different Ra=10^3, 10^4 and 10^5, respectively.

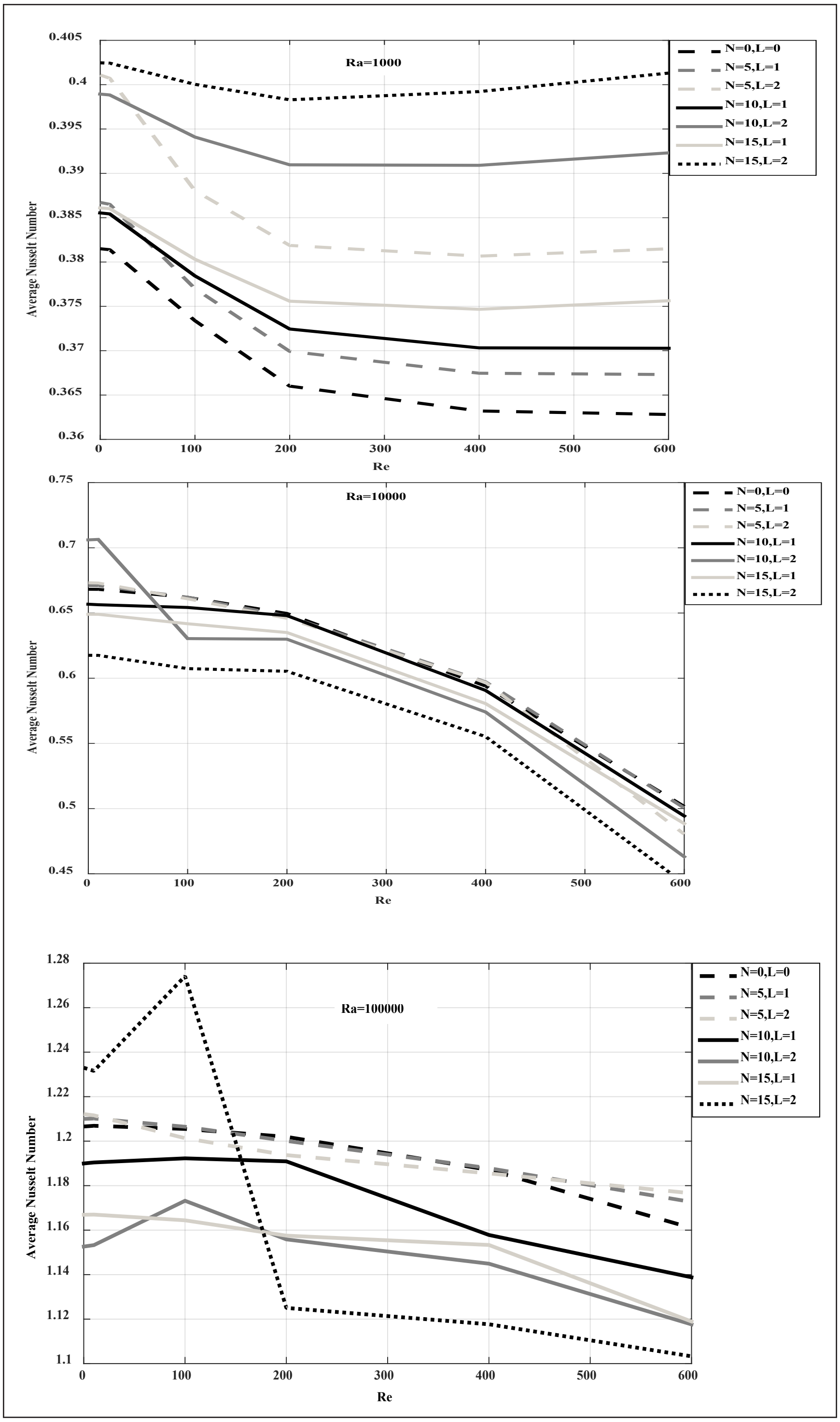


Figure 10. Average Nusselt Number as a function of amplitudes of wavy surface, undulation numbers and Reynolds numbers for Ra= 10^3, 10^4 and 10^5, respectively.

Figure 10 reveals the average Nu as a function of Re, amplitudes of the wavy surface and undulation numbers for different Ra values (Ra = 10^3, 10^4 and 10^5). At Ra = 10^3, a rapid decrease in the average Nu was observed for all cases at low Re values (0 to approximately 200). However, beyond this Re range, the average Nu was gradually decreased except when the amplitudes of the wavy surface increased to two (N = 5, N = 10 and N = 15 for L = 2) possibly due to the periodic channelling offered by the wavy surfaces of the outer cylinder, which led to increases in the heat transfer rate and the average Nu by approximately 7.3%. With an increase in Ra to 10^4, the average Nu decreased in tandem as the Re increased by approximately 2.4%. At a high Ra (Ra = 10^5), the average Nu decreased when Re increased except at (N = 10, L = 2) and (N = 15, L = 2), where the average Nu increased when Re ranged between 10 and 200 due to the rising wave amplitude.

CONCLUSIONS

The present numerical study is conducted for the mixed convection in the annular space between two concentric cylinders. The heated circular inner cylinder rotates with constant angular velocity, whereas the cold outer cylinder with a sinusoidal surface is stationary. Results show that the streamline and isotherm contours were affected by the changes in the sinusoidal surface parameters of the outer cylinder, such as amplitudes of the wavy surface and undulation numbers for different values of Re and Ra. The conclusions are as follows.

1. The isotherms and streamlines primarily depend on Ra, amplitudes of the wavy surface and undulation numbers.
2. The heat transfer from the inner cylinder increases with a rise in the surface amplitudes and undulation numbers with a fixed Re.
3. In each case, the average Nu on the surface of the inner cylinder increases with the rise of the thermal Ra and decreases throughout the flow if the rotational Re is increased.
4. The average Nu increases when the surface amplitudes and undulation numbers increase by approximately 7.3% for Ra = 10^3; it decreases as the surface amplitude and undulation numbers simultaneously increase by approximately 2.4% for Ra = 10^4 and 10^5.

ACKNOWLEDGEMENT

Authors wish to thank the University of Babylon for their help and technical support.

REFERENCES

Alam, M., Kamruzzaman, Ahsan, F., & Hasan, M. N. (2016). Mixed convection heat transfer inside a differentially heated square enclosure in presence of a rotating heat conducting cylinder. *AIP Conference Proceedings, 1754*(1), 1-8.

Alsabery, A. I., Tayebi, T., Chamkha, A. J., & Hashim, I. (2018). Effect of rotating solid cylinder on entropy generation and convective heat transfer in a wavy porous cavity heated from below. *International Communications in Heat and Mass Transfer, 95*, 197-209.

ANSYS. (2014). *ANSYS fluent theory guide (Release 16.1).* Canonsburg, Pennsylvania: ANSYS, Inc.

Bande, Y. M., & Mariah, N. A. (2014). Flat plate solar collectors and applications : a review. *Journal of Science and Technology, 22*(2), 365-385.

Boyd. R. D. (1984). Interfero metric study of natural convection heat transfer in a horizontal annulus with irregular boundaries. *Nuclear Engineering and Design, 83*(1), 105-112.

Chamkha, A. J., Hussain, S. H., & Abd-Amer, Q. R. (2011). Mixed convection heat transfer of air inside a square vented cavity with a heated horizontal square cylinder. *Numerical Heat Transfer; Part A: Applications, 59*(1), 58-79.

El-Shaarawi, M., & Sarhan, A. (1982). Combined forced-free laminar convection in the entry region of a vertical annulus with a rotating inner cylinder. *International Journal of Heat and Mass Transfer, 25*(2), 175-186.

Hussain, S. H., & Abd-Amer, Q. R. (2011). Mixed convection heat transfer flow of air inside a sinusoidal corrugated cavity with a heat-conducting horizontal circular cylinder. *Journal of Enhanced Heat Transfer, 18*(5), 433-447.

Lee. T. S. (1992). Numerical computation of fluid convection with air enclosed between the annuli of eccentric heated horizontal rotational cylinder. *Computers Fluids, 21*(3), 355-368.

Mohammed, A. (2007). Mixed convection in the entry region of a vertical annulus with a constant temperature rotating inner cylinder. *Engineering and Technology, 25*(1), 78-96.

Mozayyeni, H. R., & Rahimi, A. B. (2011). Mixed convection in cylindrical annulus with rotating outer cylinder and constant magnetic field with an effect in the radial direction. *Scientia Iranica, 19*(1), 91-105.

Rajamohan, G., Narayanaswamy, R., & Kumar, P. (2015). Experimental study on mixed convection heat transfer in a square duct with varying: inclination angles. *Journal of Science and Technology, 23*(1), 141-151.

Sheikholeslami, M., Gorji-Bandpy, M., Ganji, D. D., Soleimani, S., & Seyyedi, S. M. (2012). Natural convection of nanofluids in an enclosure between a circular and a sinusoidal cylinder in the presence of magnetic field. *International Communications in Heat and Mass Transfer, 39*(9), 1435-1443.

Sheikholeslami, M., Gorji-Bandpy, M., Pop, I., & Soleimani, S. (2013). Numerical study of natural convection between a circular enclosure and a sinusoidal cylinder using control volume based finite element method. *International Journal of Thermal Sciences, 72*, 147-158.

Sheikholeslami, M., Gorji-Bandpy, M., Ganji, D. D., & Soleimani, S. (2014). MHD natural convection in a nanofluid filled inclined enclosure with sinusoidal wall using CVFEM. *Neural Computing and Applications, 24*(3-4), 873-882.

Shekar, B. C., Vasseur, P., Robillard, L., & Nguyen, T. H. (1984). Natural convection in a heat generating fluid bounded by two horizontal concentric cylinders. *The Canadian Journal of Chemical Engineering, 62*(4), 482-489.

Shih, Y. C., Khodadadi, J. M., & Weng, K. H. (2009). Periodic fluid flow and heat transfer in a square cavity due to an insulated or isothermal rotating cylinder. *Journal of Heat Transfer, 131*(11), 1-11.

Wang, Y. F., Xu, X., Tian, T., Fan, L. W., Wang, W. L., & Yu, Z. T. (2015). Laminar mixed convection heat transfer of SiC-EG nanofluids in a triangular enclosure with a rotating inner cylinder: simulations based on the measured thermal conductivity and viscosity. *Journal of Zhejiang University-SCIENCE A, 16*(6), 478-490.

Xu, X., Yu, Z., Hu, Y., Fan, L., & Cen, K. (2010). A numerical study of laminar natural convective heat transfer around a horizontal cylinder inside a concentric air-filled triangular enclosure. *International Journal of Heat and Mass Transfer, 53*(1-3), 345-355.

Xu, X., Sun, G., Yu, Z., Hu, Y., Fan, L., & Cen, K. (2009). Numerical investigation of laminar natural convective heat transfer from a horizontal triangular cylinder to its concentric cylindrical enclosure. *International Journal of Heat and Mass Transfer, 52*(13-14), 3176-3186.

Yoo, J. (1998). Mixed convection of air between two horizontal concentric cylinders with a cooled rotating outer cylinder. *International Journal of Heat and Mass Transfer, 41*(2), 293-302.

Yuan, X., Tavakkoli, F., & Vafai, K. (2015). Analysis of natural convection in horizontal concentric annuli of varying inner shape. *Numerical Heat Transfer, Part A: Applications, 68*(11), 1155-1174.

Yu, Z. T., Fan, L. W., Hu, Y. C., & Cen, K. F. (2010). Prandtl number dependence of laminar natural convection heat transfer in a horizontal cylindrical enclosure with an inner coaxial triangular cylinder. *International Journal of Heat and Mass Transfer, 53*(7-8), 1333-1340.

Pertanika J. Sci. & Technol. 27 (3): 2015 - 2030 (2019)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Strategies to Improve Communication Management within Virtual Project Teams

Nuzul Azam Haron[1]*, Law Teik Hua[1], Salihudin Hassim[1], Fathollah Eftekhari[1], Muhammad Tahir Muhammad[1] and Aizul Nahar Harun[2]

[1]Department of Civil Engineering, Faculty of Engineering, Universiti Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia

[2]Department of Management of Technology, Malaysia-Japan International Institute of Technology, Universiti Teknologi Malaysia, Jalan Sultan Yahya Petra, 54100 Kuala Lumpur, Malaysia

ABSTRACT

Working in a virtual team presents many challenges. Communication is one of the most important challenges, especially when a virtual team includes members from different countries and background. Virtual project management enables organizations to save on resources such as cost and time. Organizations often struggle with poor communication in their geographically dispersed teams and ineffective communication have been identified as one of the main causes project failure. The aim of this study is to determine the most critical barriers to effective communication in virtual teams. A mixed method of data collection was adopted using semi structured interview with communication and construction experts, and questionnaire approach with construction companies that are G5-G7 rated. Data were analyzed using pareto and exploratory factor analysis for the development of a strategy for enhancement of communication management within virtual teams. The result shows that lack of trust and misunderstanding are the most important barrier to communication within virtual teams. However, managing communication process has been identified to have more impact on the barriers, while planning communication has the least impact. Establishing rule for response and changing focus from individual to group

ARTICLE INFO

Article history:
Received: 23 July 2018
Accepted: 21 June 2019
Published: 21 October 2019

E-mail addresses:
nuzul@upm.edu.my (Nuzul Azam Haron)
lawteik@upm.edu.my (Law Teik Hua)
hsalih@upm.edu.my (Salihudin Hassim)
heftekhar1@yahoo.com (Fathollah Eftekhari)
Mtahir1129@gmail.com (Muhammad Tahir Muhammad)
Aizulnahar.kl@utm.my (Aizul Nahar Harun)
*Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

were identified as the two most important factors required in order to sustain trust within virtual teams.

Keywords: Communication, virtual project management, virtual team

INTRODUCTION

The modern sense of project management began with the development of "Project Evaluation Review Technique" (PERT) and "Critical Path Method" (CPM) in 1958s (Stretton, 2007). Subsequently, the concept of virtual project management evolved in the mid-1990s when companies began organizing project over distance with distributed teams (Bergiel et al., 2008). Virtual project teams are emerging as an essential component for organizations to save resources such as cost and time. However, due to their structures, they create different challenges, particularly associated with communication management (Kuruppuarachchi, 2009).

However, there are many definitions of virtual project management in literature. Krill and Juell (1997), defined virtual project management as "collaborative effort towards as specific goal or accomplishment which is based on 'collective yet remote' performance". A similar definition was proposed by Rad and Levin (2003) as "working across time zone, culture, space and organizational boundaries through advanced communication and information technologies to achieve common objectives".

Virtual project teams enable organizations to quickly use human resources, experiences, capabilities, and expertise of personnel who might be in different organizations and places to provide solutions to their problems within the shortest possible time and to save cost. Communication is regarded as an integral component of success for any project team, inclusive of virtual teams to work effectively. Even though communication has been recognized as a key factor to have a successful team in a virtual environment, organizations often struggle with poor communication in their global project. A report from a study conducted by project management institute PMI (2013), revealed that 56% of every $1 billion spent on project risk was as a result of poor communication which led to failure of projects in meeting their objectives.

In the same light, Clark (2014) predicted that 70 percent of project manager did not use communication methods properly which was the main cause of the different challenges they faced while working with virtual teams. Some of these challenges are as a result of cultural diversity, different time zones, level of technology between virtual team members, lack of trust in a team and, most of all the lack of face to face communication. Communication is a very important tool for managing an engineering team and for facilitating knowledge sharing among team members (Lewkowicz et al., 2008). During the communication process, misunderstandings and errors can appear as a result of communication complexity (Hassanaly, 2006).

There are many reasons for poor communication within virtual teams, these may occur at any stage within the communication process. Some of these reasons may include the team's structure, different time zones (Aslam & Khan, 2010), language barriers (Solomon, 2010), trust (Oyeleye, 2013), and cultural differences (Shachaf, 2008). The presence of weak communication within virtual project team makes them prone to low individual commitment, role ambiguity, role overload, absenteeism, and social loafing (Jarvenpaa & Leidner, 2013).

When a team identifies barriers as the root cause of poor communication they can work to solve the issues. Unknown barriers more often have caused projects outcome be at risk. The vulnerability of these steps against obstacles is not the same in all projects. It is important to assess the existing vulnerability, and determine the degree of risk for process (Cao & Malik, 2006). Hence, due to the potential risk to the projects because of poor communication in virtual project teams, the aim of this study is to determine the critical barriers of communication management within virtual project teams and how these barriers influence on communication processes, in order to develop strategies to improve communication management in virtual project management (VPM) and achieve communication success in virtual teams.

RELATED WORK

The growing popularity of virtual teams in organization are as a result of new technological era (Walvoord et al., 2008). In addition, universal project teams are vital components of modern organizations that enable them to select the talents and expertise to innovate, solve complex problem, and save on resources (Kuruppuarachchi, 2009).

As stated by Khazanchi and Zigurs (2006), and William et al. (2010), "project management is a challenging activity in the best situations, and in the virtual environment it has become even more challenging". Manager must deal with many challenges while working with virtual teams, such as cultural diversity (Daim et al., 2012), different time zones (Aslam & Khan, 2010), technology (Hosseini & Chileshe, 2013), lack of trust in a team (Greenberg et al., 2007; Oyeleye, 2013), and, most of all, the lack of face- to- face communication (Reed & Knight, 2011; Rosen et al., 2007). Although some of these challenges have always existed for traditional project managers, as observed by Osman (2011), they are amplified several times over for virtual project managers.

Communication is an integral component of success for any project team, also for virtual teams to work effectively. Everyone has to know of outside events that will affect the team, and make sure that a problem or delay in one area is immediately communicated to those whom it may affect. As stated by Ebrahim et al. (2009), "communication is the most challenging component of project management, particularly among virtual teams, which can also make it more difficult to overcome cultural barriers." Lee-Kelley and

Sankey (2008), stated in their research that cultural differences and time zone affected communication as well as team relations on project. Similarly, cultural misunderstanding further made communication complex due to differences in language, verbal styles and nonverbal styles which influenced team effectiveness (Shachaf, 2008).

The recent study conducted by the project management institute PMI (2013), revealed that ineffective communication has negative impact on project execution. It revealed that, 56% of total project cost was at risk as a result of ineffective communication. It also showed that, ineffective communication is the primary contributor to project failure one third of the time, and had a negative impact on project success more than half the time. Similarly, Clark (2014) stated that 70% of project managers did not use communication methods properly. As such, they must deal with many challenges while working with virtual teams.

The consequence of weak communication is that the virtual project team is prone to low individual commitment, role ambiguity, role overload, absenteeism, and social loafing (Jarvenpaa & Leidner, 2013). According to PMI (2012), communication management includes three processes: plan communication, manage communication, and control & monitoring communication. The vulnerability of these steps against obstacles is not same. The only way to success in a system is to assess the existing vulnerability, and determine the degree of risk for process (Cao & Malik, 2006).

According to the recent report by cultural training service (RW3. LLC) (Solomon, 2010), 45% of employees who are working as virtual team members, claimed they had never met their virtual team groups and 30% said they only met them once a year. In the report, "the challenges of working virtual teams" was based on survey of nearly 3000 employees from multinational companies. The survey also found that virtual team members 97% did not have enough time during virtual meeting to build relationships. Similarly, 81% said that it was difficult to establish rapport and trust in virtual teams (Solomon, 2010).

The distance present in virtual teams introduces problems for effective methods of communication in task coordination as social interaction and team relationships (Anderson et al., 2007). Organizational leaders must also understand how establishing trust through social interaction affects virtual team communication (Corvello & Migliarese, 2007). The lack of effective communication, resulted by challenges, in virtual teams can make ambiguous potential understandings of objectives and complicate the communication and collaborations (Oyeleye, 2013).

Communication Management in Virtual Teams

The most satisfied virtual team members exist in teams with effective communication structures and patterns (Oyeleye, 2013). As stated by Clark et al. (2010), in order to ensure that the virtual teams meet their objectives, daily communication between a team leader and individual team member is the glue that hold a virtual team together. According to

Walvoord et al. (2008), “communication between virtual team members serves not only in the exchange of critical information among team members that are working in different places, but also to build interpersonal relationships”. Although both virtual and traditional teams share the common characteristic of communication, the difference lies in the fact that asynchronous communication is required in virtual teams (Bergiel et al., 2008).

Similarly, Bilczynska-Wojcik (2014) was of the opinion that project managers had to understand how to approach a new project with team members from various regions, countries and time zones. The researcher described that work schedules and meeting times had to be considered more than with face-to-face project teams and to reduce the impact of space and time, the project team members needed effective communication and appropriate use of the communication tools throughout the project from planning to closure. Reed and Knight (2011), also stated that while the use of virtual teams had become quite commonplace, the initiation and rapid growth of virtual project work was not accompanied by customized processes and procedures, standards, methodologies or guidelines developed specifically for the virtual environment.

Most project management practitioners instead rely upon existing traditional project risk assessment and handling methods, originally designed for co-located project teams. However, unique issues have been documented in virtual environments, including communication issues” (Lee-Kelley & Sankey, 2008), trust issues (Majchrzak et al., 2004; Powell et al., 2004), and issues with invisible team members, sometimes referred to as “deadbeats” or “freeriders” (Rubin et al., 2002). Although such issues can occur in traditional projects, these means that the problems may occur more frequently or with greater intensity when the environment is virtual. similarly, Reed and Knight (2011) stated that the lack of or inadequate communication risk factor refered to communication problems on a project that impacted project success. This included missing communication, where project team members were not adequately informed about important aspects of the project work, resulting in confusion. Inadequate communication can also occur when there is so little communication that problems result because team members do not know what to do or what is expected of them.

Communication Challenges in Virtual Project Management

There are several reasons that have made communication risk to be anticipated higher in virtual projects. However, since virtual projects have little or no face-to-face communication, team members rely on ICTs such as video conferencing, e-mails, wikis and blogs, collaboration tools and instant messaging. But in spite of their numerous advantages, virtual teams face greater communication challenges than face-to-face teams. The advancement in technology has made virtual workplace communication more prevalent for business meetings globally. Lookwood (2015), reported that most business leaders encountered more challenges in virtual communication than in face-to-face communication. However, the precise cause of these challenges is unknown due to the complexity of the business environment.

PMI's 2013 Pulse of the Profession report revealed that US$135 million is at risk for every US$1 billion spent on a project (13.5%). Further research on the importance of effective communications uncovers that a startling 56 percent (US$75 million of that US$135 million) is at risk due to ineffective communications. Contrary to this, Morgan et al. (2014) stated that the limited range of communication methods was not a major contributing factor to a team's effectiveness.

Researchers are trying to determine how virtuality impacts teams effectiveness (DeSanctis & Poole, 1997; Pauleen, 2003), focusing on a variety of success predictors (Dubé & Paré, 2001; Furst et al., 2004), such as conflict management, leadership (Kayworth & Leidner, 2002), trust (Jarvenpaa et al., 2004; Kanawattanachai & Yoo, 2007; Piccoli & Ives, 2003), communication (Maznevski & Chudoba, 2000), norm development (Majchrzak et al., 2000), boundary crossing (Espinosa et al., 2003), creativity, team size (Bradner et al., 2003), control, and technology appropriation (Majchrzak et al., 2000). The effect of virtual team composition on performance seems very important; e.g., cultural diversity influences virtual team effectiveness (Dafoulas & Macaulay, 2002).

However, literature suggests that cultural diversity is a critical predictor of effectiveness, empirical findings that support this claim are rare (Martins et al., 2004; Walsham, 2002). Therefore, it is important to gain a better understanding on the effect of cultural diversity on team effectiveness and to understand how ICT mitigates or amplifies this influence (Buckley, 1999). The concept of "trust" in virtual teams has been widely researched (Kanawattanachai & Yoo, 2007). Many dimensions of trust have been identified including cognitive trust, calculative trust and institutional trust (Erdem & Ozen, 2003). Studies have been carried out to examine trust in relation to the abilities, benevolence and integrity of team members. Trust is critical to the cooperative behavior that leads to the success of all teams, but it is especially important in virtual teams.

Two interrelated factors, diverse locations and technology-enabled communication contribute to making trust more difficult to develop in virtual teams than in traditional hierarchical relationships and on-site teams (Greenberg et al., 2007). In a nutshell, the review of related literatures reveals that there are many factors that influence the success of communication management in virtual project teams. However, missing piece of knowledge in existing literature is the evaluation of the influence of these factors on processes of communication management to identify the most vulnerable stage.

METHODOLOGY

The research process was conducted using both qualitative and quantitative methods (mixed method) in two main phases. The first phase was the qualitative study which was conducted using semi-structured interview method. A sampling frame was designed comprising two groups of academician experts in communication, and project management from three universities in Malaysia (UM, UPM, and MMU). These institutions were selected because they are the top three universities within the study area that have experts in communication, construction and virtual project management. There were 20 experts interviewed in the first phase. The purpose of the interview was to rank the important challenges of communication among virtual teams, this was conducted to identify the most critical communication challenges faced by virtual teams by conducting Pareto analysis on the items (Neuman, 2005). The second phase which was the quantitative study was conducted through the use of closed ended questionnaires to collect data from the targeted respondents. The respondents were selected from the construction contractor companies with experience of managing virtual team in project execution. The targeted population were construction companies that were registered with the Construction Industry Development Board (CIDB) Malaysia under the class G5, G6, and G7. Based on questions selected from comprehensive literature review, 118 valid responses were collected for the development of a new strategy, which focused on behaviours that affected the success of communication within virtual teams and project managers in the construction industry. The methodological framework is summarised in Figure 1 below.

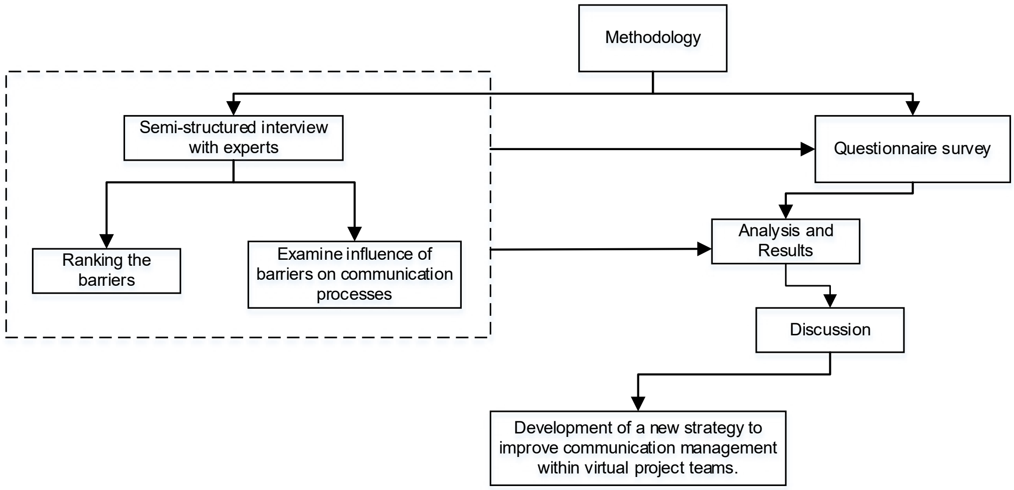


Figure 1. Methodological framework

ANALYSIS AND RESULTS

Data were collected through the use of semi-structured interview by asking participants to rank factors that are critical barriers to effective communication in virtual teams. Lack of trust, misunderstanding, level of knowledge & technology, language barriers, and multiple time zone respectively were ranked as most important communication barriers that are faced by virtual teams. This finding is in agreement with the research conducted by Greenberg et al., (2007) and Yang (2014). This shows that lack of trust is the most important challenge within virtual environments. However, the ranking in Figure 2 underscores the insignificance of some factors, such as diversity and local laws in virtual work environment.

Figure 2 shows the overall results of the ranking of communication barriers faced by virtual teams. More also, Pareto analysis was conducted on the items been ranked by the respondents. Pareto as 80% to 20% rule under the assumption that in all situations 20% of causes determine 80% of problems, this ratio is merely a convenient rule of thumb and is not nor should it be considered immutable law of nature. It should be noted that the 20% are root causes of the remaining 80%. Lack of trust and misunderstanding were selected according to the Pareto principle as the top portion of the causes that need to be addressed to resolve the majority of communication problems in virtual teams. The application of the Pareto analysis in management allows management to focus on those items that have the most impact on their project (Neuman, 2005). The result from the Pareto analysis shows that the "Lack of Trust" was observed to be the highest weight percentage of importance (13.70%), followed by "Misunderstanding" (12.63%).

The second phase of the enquiry which was quantitative, was conducted through the use of questionnaires. This was conducted in order to identify the most important factors to improve communication in virtual teams from the perspective of misunderstanding and lack of trust. The factors were analysed for reliability and validity, this was evaluated using Cronbach's coefficient alpha (α). The result of the reliability test (α) was above 0.70 as shown in Table 1. An "α" value of 0.70 or above indicates a reliable measurement

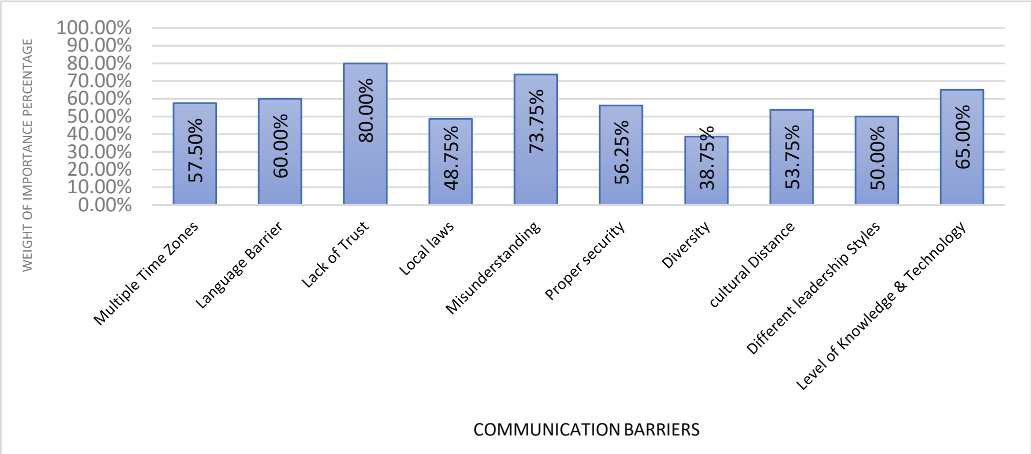


Figure 2. Importance of communication management barriers among virtual project teams.

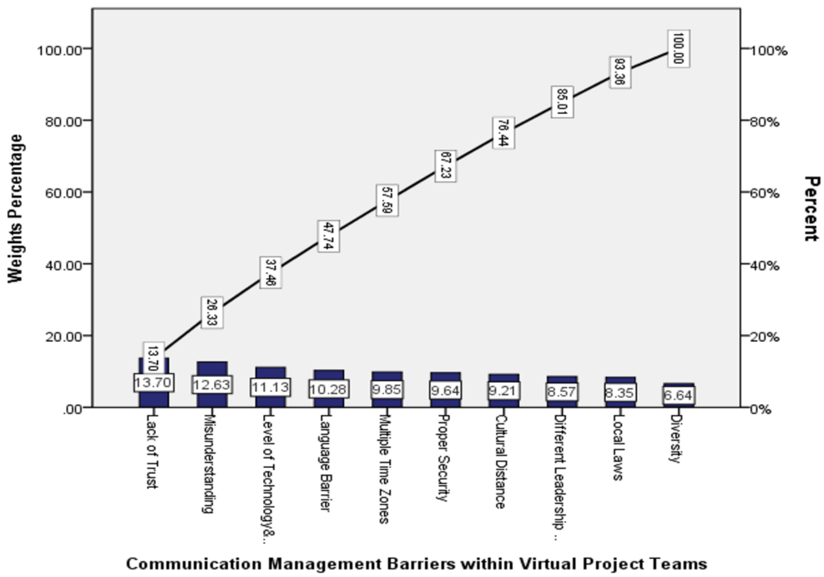


Figure 3. Pareto Analysis of Communication Management Barriers

instrument (Cronbach, 1951). Subsequently, exploratory factor analysis (EFA) was used to determine the factor structure among all variables. The principal component method was applied to evaluate the factorability of 20 items. Several well-known criteria for the factorability of correlation were used. First, Kaiser-Meyer- Olkin measure of sampling adequacy was 0.930, and Bartlett's test of sphericity was significant. The diagonals of the anti-image correlation matrix were all over 0.5, supporting the inclusion of each item in the factor analysis (Hair et al., 2006). Finally, the communalities of items were above 0.4, further confirming that each item shares some common variance with other items. As result of EFA, two components with eigenvalues greater than 1 were found (Table 2). The first component consisted of 10 items which were related to trust dimension with load point of items between 0.822 and 0.593. In the second component found in the scale, there were 10 items related to misunderstanding dimension with load point of items between 0.865 and 0.544. The four items with the highest rank from two components selected to develop a new strategy for effective communication among virtual teams (Figure 3).

In other to develop sustain trust within virtual teams, there is need to;

- Establish rules for response.
- Change focus from individual to group.
- In other to prevent misunderstanding in communication among virtual teams, there is need to;
- Use communication guideline within the virtual project teams.
- Summarizing, recapping, and identifying the next steps.

Table 1

Cronbach's Alpha Value of the Instruments

	Items	α	Cronbach's Alpha
Items to Sustained Trust	Change focus from individual to group	0.815	**0.844**
	Encourage participation of organizing activities.	0.841	
	Establish rules for response.	0.815	
	Meet face to face if practical.	0.832	
	Use the most effective method of communication.	0.851	
	Clarify tasks and processes, not just goals and roles.	0.822	
	Membership of the professional global associations.	0.806	
	Introduce team members to one another (experience & abilities).	0.846	
	Being counted on to do what the team members say they will do.	0.827	
	Share and rotate power (different members lead the team at different times)	0.840	
Items to Prevent Misunderstanding	Team member learn the different task that the team performs.	0.761	**0.792**
	Clarify individual tasks.	0.781	
	Setting and managing expectation.	0.768	
	Summarizing, recapping, and identifying the next steps.	0.756	
	Define an appropriate communication model.	0.772	
	High education to increase team's knowledge.	0.788	
	Considering all viewpoints.	0.776	
	Fact finding to identify or confirm information.	0.788	
	Selecting communication method.	0.795	
	Using communication guideline within the team.	0.750	

Table 2

Factor Analysis after extraction components using principal component method

Component	Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	11.874	59.37	59.37	7.086	35.43	35.43
2	1.452	7.258	66.628	6.24	31.20	66.63

DISCUSSIONS

Virtual teams are dispersed geographically, as such, it is important for project managers to find a clear and controllable method to communicate and cooperate. For this purpose, project managers have to select proper strategy for communication. The results of qualitative study showed that “managing communication” process had the highest impact against the barriers and “plan communication” process having the lowest impact against the barriers. Similarly, the study also found out that “establishing rule for response” and “changing focus from individual to group” were the most important factor required in order to sustain trust within virtual teams. A developed trust will also increase the pace at which virtual team builds cohesiveness.

The results of this study also reveal that team diversity shows itself in both the social and technical skill sets of the virtual team members. These are reflections of the diversities found in co-located teams, but with additional complications due to distance. Socially diverse members have varied backgrounds and life experiences and may form incorrect assessments of fellow team members. This is more pronounced within members from different cultural backgrounds that have had conflicts in the past. Statements that are acceptable by one group are often not taken in the same context as intended resulting in emotional disconnects.

Furthermore, a strategy to improve communication from perspectives of trust and misunderstanding among virtual teams is proposed. Developing new communication strategy processes starts from the determination of the objectives of communication and identifying people who are involved, affected, and interested in communication processes. A strategy provides answer to “what”, “why”, “when”, and “how”. A communication strategy is coherent narrative that illustrates a solution to an obstacle or a package of obstacles in communication management.

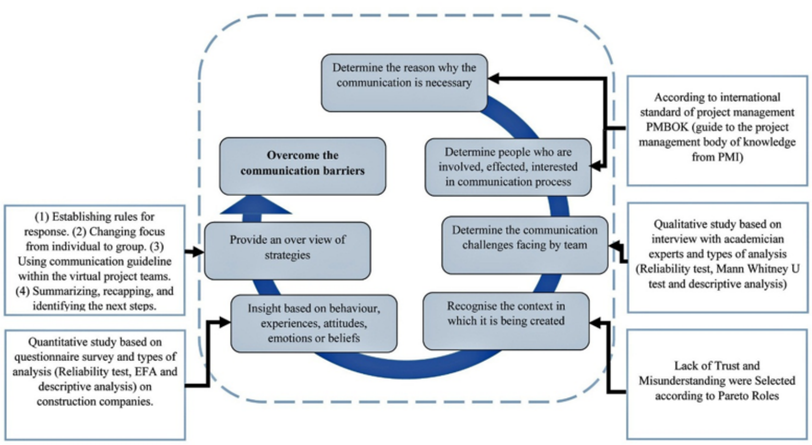


Figure 4. Process of developing communication strategy

Figure 4 above was developed from the findings of this research and it shows the proposed recommended strategies that virtual project managers should consider in communication processes (planning, managing, and monitoring) to improve their project communication programs.

In other to develop sustain trust within virtual teams, there is need to;

- Establish rules for response.
- Change focus from individual to group.
- In other to prevent misunderstanding in communication among virtual teams, there is need to;
- Use communication guideline within the virtual project teams.
- Summarizing, recapping, and identifying the next steps.

The results of this study reveal that "establishing rules for response" was the most important factor required in order to sustain trust within virtual team development stages and to improve communication management. A developed trust will also increase the pace at which the virtual team builds its cohesiveness; thus, the team members will improve their knowledge dissemination through communication. Trust is the major behavioral process associated with teams and if team member cannot communicate with each other, trust cannot be developed. Bell and Kozlowski (2002) referred to trust as the lens that helped to define the level of communication, coordination and cooperation within a virtual team. Thus, establishing rules is the first stage of developing trust between team members."

CONCLUSION AND RECOMMENDATION

It is important for organisations to understand virtual team structure and ensure that they are supported with training and resources for the success of the team. Using virtual teams in organizations is the new epoch for organizations to expand globally and to maintain their profit margins. Having a better understanding of how a virtual team should be managed is crucial to the success of an organization. It is important for organisations to resolve critical issues such as lack of trust and misunderstanding to effectively manage communication within virtual teams as well as emphasizing on the significance of trust development and reduced misunderstanding within virtual teams. More also, practitioners should focus on factors and behaviours of their different virtual teams to develop a strategy to effectively manage communication.

The findings from the study recommend the following strategies for managing communication processes;

- Establishing rules for response.
- Changing focus from individual to group.
- Using communication guideline within the virtual project teams.

- Summarizing, recapping, and identifying the next steps.

The current study can be further explored by investigating the role of communication tools to improve leadership trust within virtual teams. In addition to this, further research can be carried out to examine the impact that culture might have on trust and communication in virtual teams. Finally, further research as a follow-up to this study should be conducted in other countries paying attention to the relationship between trust and communication in other to create best practices for setting up virtual teams.

ACKNOWLEDGEMENT

Foremost, praise is to "God" the cherisher for giving me strengths, health and determination to complete this research. I would like to express my sincere gratitude and acknowledgements to my supervisory team Dr. Nuzul Azam Haron, Assoc. Prof. Dr. Salihudin Hj Hassim, and Assoc. Prof. Dr. Law Teik Hua, for their encouragement, insightful and comments. I will like to appreciate the entire management and staff of Universiti Putra Malaysia for giving me the opportunity to carry out this research in this great institution. I will like to thank my family members for supporting me always in my time of need.

REFERENCES

Aslam, S., & Khan, A. M. (2010). *Study of Effective Virtual Project Management through Emotional Intelligence, Empowerment and Leadership Style in Partly and Truly Global Projects* (Doctoral dissertation). University of Engineering and Technology Taxila, Pakistan.

Anderson, A. H., McEwan, R., Bal, J., & Carletta, J. (2007). Virtual team meetings: An analysis of communication and context. *Computers in Human Behavior, 23*(5), 2558-2580.

Bergiel, B. J., Bergiel, E. B., & Balsmeier, P. W. (2008). Nature of virtual teams: a summary of their advantages and disadvantages. *Management Research News, 31*(2), 99-110.

Bilczynska-Wojcik, A. (2014). *Communication Management within Virtual Teams in Global Projects* (Doctoral dissertation). Dublin Business School, Ireland.

Bradner, E., Mark, G., & Hertel, T. D. (2003, January 6-9). Effects of team size on participation, awareness, and technology choice in geographically distributed teams. In *Proceedings of the 36th Annual Hawaii International Conference on System Sciences* (pp. 1-10). Big Island, HI, USA.

Buckley, K. B. (1999). *A Model of Virtual Organization Effectiveness Optimizing the Outcomes of Interorganizational Project Teams* (Doctoral dissertation). George Washington University, Washington, D.C., United States.

Cao, F., & Malik, S. (2006). Vulnerability analysis and best practices for adopting IP telephony in critical infrastructure sectors. *Communications Magazine, IEEE, 44*(4), 138-145.

Clark, R. L. (2014). *Leadership Trust in Virtual Teams using Communication Tools: A Quantitative Correlational Study* (Doctoral dissertation). University of Phoenix, Tempe, Arizona.

Clark, W. R., Clark, L. A., & Crossley, K. (2010). Developing multidimensional trust without touch in virtual teams. *Marketing Management Journal, 20*(1), 177-193.

Corvello, V., & Migliarese, P. (2007). Virtual forms for the organization of production: A comparative analysis. *International Journal of Production Economics, 110*(1), 5-15.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297-334.

Dafoulas, G., & Macaulay, L. (2002). Investigating cultural differences in virtual software teams. *The Electronic Journal of Information Systems in Developing Countries, 7*(1), 1-14.

Daim, T. U., Ha, A., Reutiman, S., Hughes, B., Pathak, U., Bynum, W., & Bhatla, A. (2012). Exploring the communication breakdown in global virtual teams. *International Journal of Project Management, 30*(2), 199-212.

DeSanctis, G., & Poole, M. S. (1997). Transitions in teamwork in new organizational forms. *Advances in Group Processes, 14*(1), 57-176.

Dubé, L., & Paré, G. (2001). Global virtual teams. *Communications of the ACM, 44*(12), 71-73.

Ebrahim, N. A., Ahmed, S., & Taha, Z. (2009). Virtual teams: a literature review. *Australian Journal of Basic and Applied Sciences, 3*(3), 2653-2669.

Erdem, F., & Ozen, J. (2003). Cognitive and affective dimensions of trust in developing team performance. *Team Performance Management, 9*(5/6), 131-135.

Espinosa, J. A., Cummings, J. N., Wilson, J. M., & Pearce, B. M. (2003). Team boundary issues across multiple global firms. *Journal of Management Information Systems, 19*(4), 157-190.

Furst, S. A., Reeves, M., Rosen, B., & Blackburn, R. S. (2004). Managing the life cycle of virtual teams. *The Academy of Management Executive, 18*(2), 6-20.

Greenberg, P. S., Greenberg, R. H., & Antonucci, Y. L. (2007). Creating and sustaining trust in virtual teams. *Business Horizons, 50*(4), 325-333.

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2006). *Multivariate Data Analysis* (Vol. 6). Upper Saddle River, NJ: Pearson Prentice Hall.

Hassanaly, P. (2006). *Cooperative Systems Design: Seamless Integration of Artifacts and Conversations-Enhanced Concepts of Infrastructure for Communication* (Vol. 137). Amsterdam, Netherlands: IOS Press.

Hosseini, M. R., & Chileshe, N. (2013). Global Virtual Engineering Teams (GVETs): A fertile ground for research in Australian construction projects context. *International Journal of Project Management, 31*(8), 1101-1117.

Jarvenpaa, S. L., & Leidner, D. E. (2013). Communication and trust in global virtual teams. *Journal of Computer-Mediated Communication, 3*(4), 0-0.

Jarvenpaa, S. L., Shaw, T. R., & Staples, D. S. (2004). Toward contextualized theories of trust: The role of trust in global virtual teams. *Information Systems Research, 15*(3), 250-267.

Kanawattanachai, P., & Yoo, Y. (2007). The impact of knowledge coordination on virtual team performance over time. *MIS quarterly, 31*(4), 783-808.

Kayworth, T. R., & Leidner, D. E. (2002). Leadership effectiveness in global virtual teams. *Journal of Management Information Systems, 18*(3), 7-40.

Khazanchi, D., & Zigurs, I. (2006). Patterns for effective management of virtual projects: Theory and evidence. *International Journal of e-Collaboration (IJeC), 2*(3), 25-49.

Krill, T., & Juell, P. (1997, March). Virtual project management. In *Proceedings of the Small College Computing Symposium (SCCS'97)*. North Dakota State University, Fargo, North Dakota.

Kuruppuarachchi, P. R. (2009). Virtual team concepts in projects: a case study. *Project Management Journal, 40*(2), 19-33.

Lee-Kelley, L., & Sankey, T. (2008). Global virtual teams for value creation and project success: A case study. *International Journal of Project Management, 26*(1), 51-62.

Lewkowicz, M., Wijnhoven, F., & Draghici, A. (2008). Misunderstandings in global virtual engineering teams: definitions, causes, and guidelines for knowledge sharing and interaction. In A. Bernard & S. Tichkiewitch (Eds.), *Methods and Tools for Effective Knowledge Life-Cycle-Management* (pp. 145-157). Berlin, Heidelberg: Springer.

Lockwood, J. (2015). Virtual team management: what is causing communication breakdown? *Language and Intercultural Communication*, *15*(1), 125-140.

Majchrzak, A., Malhotra, A., Stamps, J., & Lipnack, J. (2004). Can absence make a team grow stronger? *Harvard Business Review, 82*(5), 131-137.

Majchrzak, A., Rice, R. E., Malhotra, A., King, N., & Ba, S. (2000). Technology adaptation: The case of a computer-supported inter-organizational virtual team. *MIS quarterly, 24*(4), 569-600.

Martins, L. L., Gilson, L. L., & Maynard, M. T. (2004). Virtual teams: What do we know and where do we go from here? *Journal of management, 30*(6), 805-835.

Maznevski, M. L., & Chudoba, K. M. (2000). Bridging space over time: Global virtual team dynamics and effectiveness. *Organization Science, 11*(5), 473-492.

Morgan, L., Paucar-Caceres, A., & Wright, G. (2014). Leading effective global virtual teams: The consequences of methods of communication. *Systemic Practice and Action Research*, *27*(6), 607-624.

Neuman, W. L. (2005). *Social Research Methods: Quantitative And Qualitative Approaches* (Vol. 13). Boston, Massachusetts: Allyn and Bacon.

Osman, H. (2011). *The Virtual Project Manager: Seven Best Practices for Effective Communication*. PMI Virtual Library. Retrieved October 10, 2014, from https://www.projectmanagement.com/articles/284882/The-Virtual-Project-Manager--Seven-Best-Practices-for-Effective-Communication

Oyeleye, O. O. (2013). *Trust and Virtual Teams: The Influence of Leadership Behavior, Communication, and Gender Behavior of Virtual Team Leaders on the Development of Trust in Virtual Teams* (Doctoral dissertation). University of Maryland University College, Adelphi, Maryland.

Pauleen, D. J. (2003). Lessons learned crossing boundaries in an ICT-supported distributed team. *Journal of Global Information Management (JGIM), 11*(4), 1-19.

Piccoli, G., & Ives, B. (2003). Trust and the unintended effects of behavior control in virtual teams. *MIS quarterly, 27*(3), 365-395.

PMI. (2012). *Guide to the Project Management Body of Knowledge Project Management Institute*. Newton Square, Pennsylvania: Project Management Institute.

PMI. (2013). *The essential role of communications.* Retrieved October 10, 2014, from http://www.pmi.org/~/media/PDF/Business-Solutions/The-High-Cost-Low-Performance-The-Essential-Role-of-Communications.ashx

Powell, A., Piccoli, G., & Ives, B. (2004). Virtual teams: a review of current literature and directions for future research. *ACM Sigmis Database, 35*(1), 6-36.

Rad, P. F., & Levin, G. (2003). *Achieving Project Management Success using Virtual Teams.* Boca Raton, Florida: J. Ross Publishing.

Reed, A. H., & Knight, L. V. (2011). Major virtual project risk factors. *Journal of Information Technology Management, 22*(4), 1-12.

Rosen, B., Furst, S., & Blackburn, R. (2007). Overcoming barriers to knowledge sharing in virtual teams. *Organizational Dynamics, 36*(3), 259-273.

Rubin, R. S., Bommer, W. H., & Baldwin, T. T. (2002). Using extracurricular activity as an indicator of interpersonal skill: Prudent evaluation or recruiting malpractice? *Human Resource Management, 41*(4), 441-454.

Shachaf, P. (2008). Cultural diversity and information and communication technology impacts on global virtual teams: An exploratory study. *Information and Management, 45*(2), 131-142.

Solomon, C. (2010). *The Challenges of Working in Virtual Teams: Virtual Teams Survey Report 2010.* New York, NY: RW3 CultureWizard.

Stretton, A. (2007). A short history of modern project management. *PM World Today, 9*(10), 1-18.

Walsham, G. (2002). Cross-cultural software production and use: a structurational analysis. *MIS quarterly, 26*(4), 359-380.

Walvoord, A. A., Redden, E. R., Elliott, L. R., & Coovert, M. D. (2008). Empowering followers in virtual teams: Guiding principles from theory and practice. *Computers in Human Behavior, 24*(5), 1884-1906.

William, Dow, P., & Taylor, B. (2010). *Project Management Communications Bible* (Vol. 574). Indianapolis, Indiana: John Wiley & Sons.

Yang, I. (2014). What makes an effective team? The role of trust (dis) confirmation in team development. *European Management Journal, 32*(6), 858-869.

Pertanika J. Sci. & Technol. 27 (4): 2031 - 2047 (2019)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Assessment of External Risk Factors on Construction Project Schedule Using Risk Importance Index (RII)

Saiful Husin*, Abdullah, Medyan Riza and Mochammad Afifuddin

Department of Civil Engineering, Faculty of Civil Engineering, Universitas Syiah Kuala, 23111 Aceh, Indonesia

ABSTRACT

External risk factors influence the project objectives, especially project scheduling. Some risks may occur frequently but have low impact, while others may have both high frequency and high severity. The linkage between risk frequency and impact needs to be assessed to understand the significance of a risk variable. This study was aimed to examine the significance of risk related to the probability of occurrence and the magnitude of impact on the project schedule. The study focused on external risk factors (sociopolitical, government policy, natural disasters, and monetary). Survey questionnaires were sent to 20 targeted contractor companies in the area of Aceh Province, Indonesia with a response rate of (60%). Using Risk Importance Index (RII), the significant risk variables that hindered the achievement of project achievement in Period I have been identified as follows: cultural conditions and local customs near project site (K4), social issues/surrounding environment (K5), increased fuel prices (L4), uncertain weather conditions (M1). In addition, the emergence of L4 variable in Period I was due to temporary security conditions, while in Period II due to the conditions of supply and demand. The results of this study are intended to contribute to the application of risk to practitioners and governments.

Keywords: Construction project, contractor, external risk, project schedule, risk importance

ARTICLE INFO

Article history:
Received: 12 September 2018
Accepted: 17 May 2019
Published: 21 October 2019

E-mail addresses:
saifulhusin@unsyiah.ac.id (Saiful Husin)
abdullahmahmud@unsyiah.ac.id (Abdullah)
medyan_riza@unsyiah.ac.id (Medyan Riza)
m.afifuddin@unsyiah.ac.id (Mochammad Afifuddin)
*Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

INTRODUCTION

During the implementation of a construction project, project management must minimize risks to achieve project objectives and avoid negative impacts, such as cost overruns, time delays, and quality deterioration. The risk is minimized by selecting corrective actions using decision-making based on an analysis of the various risk factors. In a construction project, the inherent risk factor is uncertainty, which is related to the risk of an event. The risk is a consequence (or outcome) of activity as determined by human judgment (Aven & Renn, 2009). Risk as the concept of an opportunity is defined as unexpected events that may occur with various consequences resulting in delays or even failures in a project (Gray & dan Larson, 2000). Risk can also be seen as an event that, if it occurs, will impact the project outcome (Clayton, 2011) by causing a failure to achieve the planned goals regarding cost, time, and project performance (Kerzner, 2009).

The uncertainties are potentially minimized by performing a risk analysis to identify the possibility of occurrence of risk factors in the project implementation. Construction project risk can arise from various sources namely internal factors or external factors (Akintoye & MacLeod, 1997; Ward & Chapman, 2003). Internal risk factors can be derived from project resources factors (Husin et al., 2018; Husin et al., 2017; Zhao et al., 2013), financial factors (Fachrurrazi et al., 2018; Farrel, 2003), project managerial and operational factors (Latham & Braun, 2008), design and contract factors (Moazzami et al., 2011), while external risk factors come from socio-political (Khodeir & Mohamed, 2015), government policy (Banaitiene & Banaitis, 2012; Pheng et al., 2008), natural disasters, and monetary (Mubarak et al., 2017).

The appearance of risk to a construction project can be associated with the condition of a region related to the potential of disasters (Moe & Pathranarakul, 2006; Christoplos et al., 2001) or vulnerabilities (Zhang, 2007; Fidan et al., 2011). The vulnerability conditions can be shown by learning from the situation in the Aceh-Indonesia. The province experienced political upheavals and tensions with the central government over the past 15 years and experiencing the earthquake and tsunami disaster. During the years, Aceh Province has recently experienced three important periods (Zeccola, 2011). The periods defined as the Period I, the period of political and military conflict; Period II, the period of rehabilitation and reconstruction post-tsunami disaster; and Period III, the period of post-rehabilitation and reconstruction. These periods have different characteristics between one another.

The previous studies indicated that risks related to project resources and external factors tend to have the high frequency of occurrence and impact to project cost (Husin et al., 2017; Mubarak et al., 2017). Discussion of risk related to project completion time is required in further study to understand the significance of risk factors and variables. The appearance of risk variables to the project completion time often results in delays and possibly affecting project postpone. Based on these conditions, this study aimed to assess the potential risks that might affect project schedule completion. The assessment focused on external factors consisting of socio-political factors, government policies, natural disasters, and monetary factors. This study examined the significance of risk importance related to the probability of occurrence and impact magnitude to the achievement of the project due date.

METHODS

Data Collection

This study began with primary data collection using a questionnaire instrument. Questionnaires were prepared to provide three types of information: (1) information related to the characteristics of respondents, (2) information related to the frequency of risk factor occurrence, and (3) information related to the impact on the completion of construction. Four external risk factors were the focus of this research: socio-political (seven variables), government policy (five variables), natural disaster (nine variables), and monetary (five variables) as mentioned in Table 1. To analyse the risk variables, quantitative and qualitative mix methods were used.

The information collected was related to the condition of the study area as described by Zeccola (2011). The periods and year ranges were divided into Period I (the year 2000-2004), Period II (the year 2005-2009), and Period III (the year 2010-2015). This study involved a number of respondents from 15 local contractor companies with large qualifications in Aceh Province. The respondents were personnel in construction companies with positions of directors, managers, and senior engineers.

Table 1

List of project external risk factors and variables

Risk Factor	Code	Variable
Social politics	K1	Riot
	K2	Sabotage of facilities or materials
	K3	Demonstration at the project site
	K4	Cultural conditions and local customs near the project site
	K5	Social issues/surrounding environment
	K6	Conflict with project stakeholders
	K7	Religious holidays or other holidays
Government policy	L1	Government policy changes or revisions that halt the project
	L2	Changes in government regulations
	L3	Late permissions and licenses (pre-implementation)
	L4	Increased fuel prices
	L5	Project cancellation by government

Table 1 (Continued)

Risk Factor	Code	Variable
Natural disasters	M1	Uncertain weather conditions
	M2	The uncertainty of field conditions
	M3	War
	M4	Revolution
	M5	Fire
	M6	Environmental pollution
	M7	Disease epidemic
	M8	Flooding
	M9	Earthquake
Monetary	N1	Monetary instability
	N2	Fluctuations in bank loan interest rates
	N3	Long-term currency inflation, deflation, and devaluation
	N4	Short-term currency fluctuations
	N5	Economic crisis

Sample Adequacy Test

The theory that is widely used in determining sample size is the Slovin formula (Ariola, 2006; Eduardus & Hamsa 2013):

$$n = \frac{N}{1 + N\,(e)^2} \tag{1}$$

Where n is the number of samples, N is the population, and e is the accuracy of sampling errors (0.1 or 10% for the large population; and 0.2 or 20% for the small population).

Questionnaire Response Tests

The questionnaire response validity was tested using the Pearson product moment correlation (Equations 2), and its reliability was tested using the *Cronbach-alpha* (*C-alpha*) method (Smith, 2015). The decision criteria for determining the validity of the item was based on the following conditions:

If $r_{stat} > r_{sig}$, then the question item is declared valid.

If $r_{stat} < r_{sig}$, then the question item is declared invalid.

$$r = \frac{n\,(\sum xy) - (\sum x \,.\, \sum y)}{\sqrt{\{n \sum x^2 - (\sum x)^2\}\{n \sum y^2 - (\sum y)^2\}}}\ , \tag{2}$$

where r is a correlation coefficient, Σx is the total item score, Σy is the total score, and n is the number of respondents.

The reliability test *(r)* is analyzed with Equations 3, 4, and 5 (Bland & Altman, 1997), and the feasibility indicator coefficient uses of *C-alpha* ≥ 0.6 with the significance level (α) of 5%.

$$r = \frac{k}{(k-1)}\left[1 - \frac{{\sigma_b}^2}{{\sigma_1}^2}\right], \tag{3}$$

$${\sigma_b}^2 = \frac{JK_i}{n} - \frac{JK_S}{n^2}, \tag{4}$$

$${\sigma_1}^2 = \frac{\sum {x_t}^2}{n} - \frac{(\sum x_t)^2}{n^2}, \tag{5}$$

Where r is the instrument reliability, k is the number of questions, ${\sigma_b}^2$ is the item variance, ${\sigma_1}^2$ is the total variance, Σx_t is the total number of respondent answers, $\sum {x_t}^2$ is the squared of total number respondent answers, Jk_i is the sum of squares for whole items, and Jk_s is the sum of squares for subjects.

Analysis of Frequency Index (*FI*)

The frequency index (*FI*) is used to measure the risk variable frequency. The *FI* assessment was classified according to the five-level Likert scale as shown in Table 2 and analyzed by using Equation 6 (Majid & McCaffer, 1997).

$$FI = \frac{\sum_{i=1}^{5} a_i n_i}{5N}, \tag{6}$$

Where i is the index scale of the response, a_i is the weight of the i-th response, n_i is the frequency of the variable in all responses, and N is the total number of respondents.

Analysis of Severity Index (*SI*)

This study used a severity index (*SI*) to represent the severity of risk impact for all observed variables. The *SI* assessment levels were classified according to the five-level Likert scale, as shown in Table 3. The formula to calculate the *SI* (Majid & McCaffer, 1997) is in Equation 7.

$$SI = \frac{\sum_{i=1}^{5} a_i n_i}{5N}, \tag{7}$$

Where i is the index scale of the response, a_i is the weight of the i-th response, n_i is the impact of the variable occurrence in the total response, and N is the total number of respondents.

Table 2

The *FI Assessment Criteria and Scales*

Subjective Frequency	Likert Scale	Assessment Scale
Very rarely	1	$0 \leq FI \leq 0.125$
Rarely	2	$0.125 < FI \leq 0.375$

Table 3

SI Assessment Criteria and Scales

Subjective Frequency	Likert Scale	Assessment Scale
Very low	1	$0 \leq SI \leq 0.125$
Low	2	$0.125 < SI \leq 0.375$
Medium	3	$0.375 < SI \leq 0.625$
High	4	$0.625 < SI \leq 0.875$
Very high	5	$0.875 < SI \leq 1.000$

Analysis of Risk Importance Index (*RII*)

The risk importance index (*RII*) is a method for measuring the importance of risk based on its frequency and severity, as represented by Equation 8. The importance of risk is analyzed for each external risk variable.

$$RII = FI \ X \ SI, \tag{8}$$

The *RII* risk scale is mapped onto a risk scale matrix (Figure 1). These values are compared with the scale to identify qualitative risk cells. Risk assessment is based on the matrix scale of the value of the risk importance index, where the lowest is 0 and the highest is 1. The scale is divided into five categories, where the range of each category is low (L), medium (M), high (H), significant (S), and extreme (E).

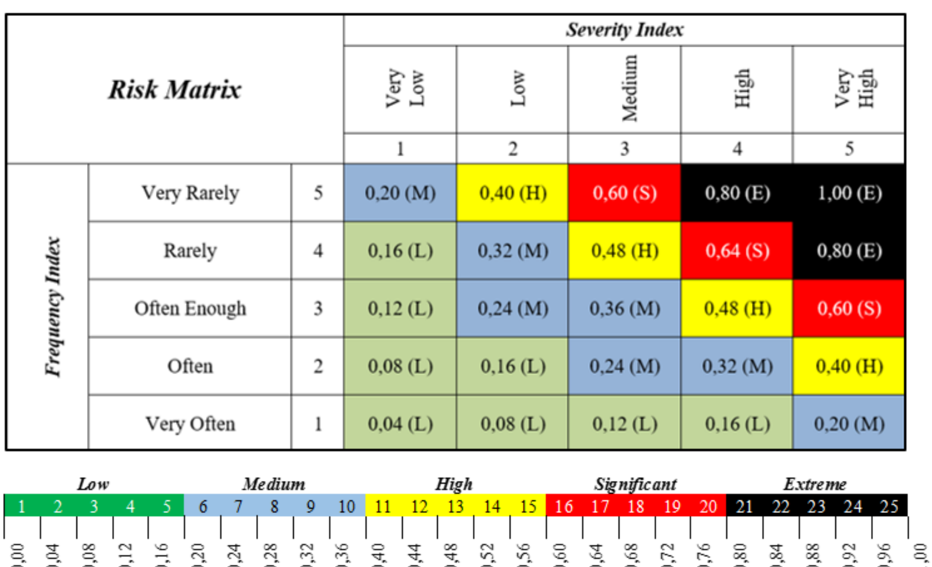


Figure 1. Risk matrix and classification of *RII*

RESULTS AND DISCUSSION

Characteristics of Respondents

The characteristics of respondents involved in gathering information in this study are presented in Table 4. Personnel involved as respondents were representatives of 15 large qualified contracting companies domiciled in Aceh Province. Most of the respondents in this study are directors and company managers with work experience of more than seven years. The number of projects that have been handled by the contractor in the conflict period is less than the next two periods. The projects handled are generally road and bridge projects with a value of 10 billion to 50 billion in Indonesian Rupiah (IDR).

Table 4
Characteristics of respondents

	Characteristic	Category Measurement	Amount	(%)
Personnel Profile	Position Respondents	Director	5	33,3
		Manager	7	46,67
		Others	3	20,00
	Respondents Experience	>2-4 years	1	6,67
		>4-7 years	1	6,67
		>7 years	13	86,67
Companies Profile	Total of contractors based on the number of projects handled	Period I		
		1-3 projects	3	20.00
		>3-6 projects	5	33.33
		>6-10 projects	4	30.00
		>10 projects	3	40.00
		Period II		
		1-3 projects	1	6.67
		>3-6 projects	5	33.33
		>6-10 projects	3	20.00
		>10 projects	6	40.00
		Period III		
		1-3 Projects	2	13.33
		>3-6 projects	2	13.33
		>6-10 projects	5	33.33
		>10 projects	6	40.00

Table 4 (Continued)

Characteristic	Category Measurement	Amount	(%)
Types of projects handled	Building	11	73.33
	Roads and bridges	14	93.33
	Water constructions	9	60.00
Average of Contract Price Yearly (in IDR)	<10 Billion	2	13.33
	10 Billion-50 Billion	8	53.33
	>50 Billion	2	13.33

The Result of Sampling Adequacy, Validity and Reliability Tests

The population of the study was restricted to local contractor companies with large qualifications and had been established before the year 2000 in Aceh Province. According to the data obtained from the Construction Services Development Board (2016), the total companies are 20 companies. This study applied population data from 20 companies and 20% sampling errors for small populations. Based on these data, this study sets a minimum sample size of 12 companies. With a total sample of 15 respondents, the adequacy of the data could be stated to have been fulfilled.

The validity test in this study based on instrument testing result of $r \geq r_{sig}$ is a valid instrument, while the instrument with $r \leq r_{sig}$ is invalid. The r_{sig} of 0,514 referred to Pearson product-moment values for a significance level of 95% and the total sample number (*n*) of 15 (Husin et al., 2018). Thus, the result of the validity test for both the frequency and severity of 24 risk variables are respectively summarised in Tables 5 and 6.

The result of the reliability test indicates that the C-alpha values for all variables of labor risk factors, materials, and project equipment for data frequency and severity were higher than 0.6. The frequency and severity data for validity test results are summarised in Tables 7 and 8, respectively.

Risk Assessment

When analyzed based on risk variables, three variables stood out on the frequency scale, namely K1 (riot), K4 (cultural conditions and local customs), and K5 (social issues) (Table 9). The K5 variable was a consistent variable appearing with frequent results for all three study periods, while the K1 and K4 variables only appeared in period I. Among government policy factors, only one variable – L4 (fuel prices) – appeared with a frequency of "often," and it appeared in all three phases of the study. Among natural disaster risk factors, three variables appeared with a frequency of "often": M1 (weather conditions), M3 (war), and M8 (flooding). The M1 variable appeared "frequently" in all three periods, while the M3 and M8 variables only appeared in one study period each, periods I and III, respectively.

Among the monetary risk factors, only one variable – N5 (economic crisis) – appeared with a frequency of "often" and only in Period I.

Table 5
Results of validity test of frequency data

Variable Code	The range of r_{stat} Value per Period			Result
	Period I	Period II	Period III	
K1; K2; K3; K4; K5; K6; K7	0,593; 0,743; 0,894; 0,542; 0,557; 0,825; 0,541	0,782; 0,760; 0,761; 0,460; 0,613; 0,761; 0,549	0,841; 0,793; 0,862; 0,548; 0,706; 0,794 0,540	Valid
L1; L2; L3; L4; L5	0,933; 0,921; 0,854; 0,551; 0,535	0,946; 0,963; 0,689; 0,560; 0,570	0,909; 0,933; 0,801; 0,525; 0,560	Valid
M1; M2; M3; M4; M5; M6; M7; M8; M9	0,528; 0,861; 0,578; 0,826; 0,878; 0,739; 0,743; 0,667; 0,909	0,530; 0,793; 0,700; 0,833; 0,774; 0,828; 0,831; 0,516; 0,561	0,578; 0,701; 0,759; 0,806; 0,856; 0,848; 0,864; 0,645; 0,861	Valid
N1; N2; N3; N4; N5	0,531; 0,936; 0,917; 0,859; 0,558	0,656; 0,880; 0,887; 0,737; 0,689	0,655; 0,922; 0,903; 0,760; 0,680	Valid

Table 6
Results of validity test of severity data

Variable Code	Range of r_{stat} Value per Period			Result
	Period I	Period II	Period III	
K1; K2; K3; K4; K5; K6; K7	0,714; 0,820; 0,816; 0,588; 0,720; 0,800; 0,763	0,816; 0,859; 0,790; 0,559; 0,869; 0,836; 0,820	0,832; 0,831; 0,935; 0,857; 0,887; 0,974; 0,785	Valid
L1; L2; L3; L4; L5	0,870; 0,822; 0,850; 0,802; 0,763	0,761; 0,849; 0,773; 0,870; 0,823	0,916; 0,881; 0,903; 0,843; 0,936	Valid
M1; M2; M3; M4; M5; M6; M7; M8; M9	0,591; 0,937; 0,628; 0,865; 0,908; 0,913; 0,922; 0,658; 0,907	0,568; 0,926; 0,841; 0,896; 0,917; 0,918; 0,946; 0,630; 0,745	0,535; 0,911; 0,853; 0,929; 0,911; 0,945; 0,938; 0,524; 0,908	Valid
N1; N2; N3; N4; N5	0,661; 0,916; 0,943; 0,800; 0,530	0,983; 0,899; 0,944; 0,926; 0,895	0,952; 0,961; 0,948; 0,961; 0,908	Valid

Table 7
Results of reliability test of frequency data

Factor	Questionnaire Reliability Score			Result
	Period I	Period II	Period III	
Social politics	0.83	0.84	0.85	Reliable
Government policy	0.80	0.79	0.77	Reliable
Natural disasters	0.88	0.86	0.89	Reliable
Monetary	0.76	0.79	0.79	Reliable

Table 8
Results of reliability test of severity data

Factor	Questionnaire Reliability Score			Result
	Period I	Period II	Period III	
Social politics	0.78	0.82	0.87	Reliable
Government policy	0.75	0.75	0.81	Reliable
Natural disasters	0.87	0.88	0.89	Reliable
Monetary	0.78	0.81	0.82	Reliable

Table 9
Risk Result for Frequency Index (FI)

Var.	Period I		Period II		Period III	
	FI	Scale	FI	Scale	FI	Scale
K1	0,640	Often	0,533	Somewhat Often	0,347	Rarely
K2	0,360	Rarely	0,333	Rarely	0,307	Rarely
K3	0,360	Rarely	0,293	Rarely	0,360	Rarely
K4	0,720	Often	0,587	Somewhat Often	0,560	Somewhat Often
K5	0,853	Often	0,667	Often	0,627	Often
K6	0,307	Rarely	0,347	Rarely	0,347	Rarely
K7	0,587	Somewhat Often	0,453	Somewhat Often	0,440	Somewhat Often
L1	0,587	Somewhat Often	0,613	Somewhat Often	0,507	Somewhat Often
L2	0,480	Somewhat Often	0,507	Somewhat Often	0,440	Somewhat Often
L3	0,547	Somewhat Often	0,493	Somewhat Often	0,440	Somewhat Often
L4	0,840	Often	0,813	Often	0,627	Often
L5	0,347	Rarely	0,400	Somewhat Often	0,360	Rarely

Table 9 (Continued)

Var.	Period I		Period II		Period III	
	FI	Scale	FI	Scale	FI	Scale
M1	0,800	Often	0,693	Often	0,653	Often
M2	0,413	Somewhat Often	0,373	Rarely	0,347	Rarely
M3	0,627	Often	0,280	Rarely	0,240	Rarely
M4	0,320	Rarely	0,267	Rarely	0,267	Rarely
M5	0,333	Rarely	0,320	Rarely	0,347	Rarely
M6	0,360	Rarely	0,280	Rarely	0,280	Rarely
M7	0,347	Rarely	0,293	Rarely	0,293	Rarely
M8	0,400	Somewhat Often	0,560	Somewhat Often	0,627	Often
M9	0,320	Rarely	0,547	Somewhat Often	0,307	Rarely
N1	0,360	Rarely	0,307	Rarely	0,280	Rarely
N2	0,427	Somewhat Often	0,440	Somewhat Often	0,400	Somewhat Often
N3	0,467	Somewhat Often	0,427	Somewhat Often	0,413	Somewhat Often
N4	0,427	Somewhat Often	0,440	Somewhat Often	0,400	Somewhat Often
N5	0,680	Often	0,373	Rarely	0,480	Somewhat Often

Analysis of the SI for external risk factors showed that no variables had a severity scale of "often," but several had a medium severity scale (Table 10). From the socio-political risk factors, four variables had a medium scale: K1 (riot), K4 (cultural conditions and local customs), K5 (social issues), and K7 (holidays). Of these variables, the variables that scored "medium" in all three study periods were K4 and K5. The variable K1 appeared consistently only in periods I and II, while K7 variable appeared only in Period I. Among the government policy risk factors, variables with consistently moderate severity scales were L1 (disruption government policy changes), L3 (license delays), and L4 (fuel prices) during all three study periods. Variable L3 appeared consistently in periods I and II, whereas L1 appeared only in Period I. Among natural disaster risk factors, two variables had a medium risk, M8 (flooding) and M9 (earthquakes). Variable M8 consistently appeared in all three study periods, while M9 only appeared as a medium risk during Period II. Among the monetary risk factors, the variable with medium risk was N4 (currency fluctuations), but it appeared only during Period I.

Table 10
Risk Results for Severity Index (SI)

Var.	Period I		Period II		Period III	
	SI	Scale	SI	Scale	SI	Scale
K1	0,410	Medium	0,410	Medium	0,37	Low
K2	0,310	Low	0,290	Low	0,29	Low
K3	0,270	Low	0,280	Low	0,32	Low
K4	0,520	Medium	0,410	Medium	0,4	Medium
K5	0,550	Medium	0,390	Medium	0,39	Medium
K6	0,320	Low	0,280	Low	0,37	Low
K7	0,410	Medium	0,360	Low	0,32	Low
L1	0,390	Medium	0,320	Low	0,37	Low
L2	0,350	Low	0,360	Low	0,35	Low
L3	0,400	Medium	0,410	Medium	0,36	Low
L4	0,470	Medium	0,450	Medium	0,41	Medium
L5	0,290	Low	0,330	Low	0,32	Low
M1	0,510	Medium	0,430	Medium	0,44	Medium
M2	0,330	Low	0,330	Low	0,29	Low
M3	0,350	Low	0,290	Low	0,28	Low
M4	0,280	Low	0,280	Low	0,29	Low
M5	0,310	Low	0,320	Low	0,32	Low
M6	0,290	Low	0,310	Low	0,31	Low
M7	0,320	Low	0,310	Low	0,32	Low
M8	0,410	Medium	0,450	Medium	0,44	Medium
M9	0,310	Low	0,400	Medium	0,31	Low
N1	0,320	Low	0,290	Low	0,29	Low
N2	0,310	Low	0,290	Low	0,35	Low
N3	0,320	Low	0,280	Low	0,33	Low
N4	0,400	Medium	0,310	Low	0,33	Low
N5	0,370	Low	0,330	Low	0,28	Low

Social Politics Risk Assessment

The calculation of average RII for each period revealed that the socio-political risk factor with the highest RII during all three periods was variable K5 (social issues) (Table 11). The risk of K5 over time tended to decrease successively from Period I (high, 0.469) to Period II (medium, 0.260) to Period III (medium, 0.245).

Table 11
Results of Risk Importance Index (RII) for Social Politics Factors

Variable Code	Period I		Period II		Period III	
	RII	Assessment	RII	Assessment	RII	Assessment
K1	0.262	Medium	0.219	Medium	0.128	Low
K2	0.112	Low	0.097	Low	0.089	Low
K3	0.097	Low	0.082	Low	0.115	Low
K4	0.374	High	0.241	Medium	0.224	Medium
K5	0.469	High	0.260	Medium	0.245	Medium
K6	0.098	Low	0.097	Low	0.128	Low
K7	0.241	Medium	0.163	Medium	0.141	Low

The high risk of variable K4 on Period I was caused by rework that was risky to project delay. This condition is in line with Abeku et al. (2016) who stated that the rework was an unwanted or undesirable of the contractor for works schedule, wastages. The K4 variable in “high” scale had decreased to become “medium” in the following period, due to a lot of training and workshops conducted by the government and NGOs in period II to increase labor capacity in terms of competence and discipline of work. This condition indirectly enhanced culture and customs in Aceh.

The K5 variable had the same pattern as the K4 variable. Variable K5 is a variable that can be caused by the conditions of conflict in Period I and disasters in Period II. However, in Period II there was a decrease caused by the recovery program from the Aceh-Nias NGO and Rehabilitation and Reconstruction Agency (BRR). As stated by Zeccola (2011), NGOs have a significant role in building the character of humanitarianism, especially when natural and political disasters collide, and resolving social problems in the Aceh Province.

Government Policy Risk Assessment

The increase in fuel prices (L4) in Aceh Province in both Period I and Period II was local in the context of the Aceh region (Table 12). The variable of increase in fuel prices (L4) occured in the scale of “High” on Period I and Period II, but then it decreased to be “Medium” on the Period III. In Period I, the high risk of time from the L4 variable was caused by scarcity of fuel that occurred due to problems with supply and transportation to the location of the conflict area in Aceh. While in Period II, it was caused by high demand during the post-tsunami rehabilitation and reconstruction phase. The decline in fuel prices to normal in Period III was automatically caused by a decline in demand for fuel in Period III.

Table 12

Results of RII for Government Policy Factors

Variable Code	Period I		Period II		Period III	
	RII	Assessment	RII	Assessment	RII	Assessment
L1	0.229	Medium	0.196	Medium	0.188	Medium
L2	0.168	Medium	0.183	Medium	0.154	Low
L3	0.219	Medium	0.202	Medium	0.158	Low
L4	0.395	High	0.366	High	0.257	Medium
L5	0.101	Low	0.132	Low	0.115	Low

Natural Disasters Risk Assessment

Variable M1 (weather) was a time risk variable that occured in Period I ("High"), the high risk on the project schedule caused by the M1 variable is not related to the condition of the conflict, but only coincidentally in Period 1 it was caused by global weather as well could occured in other regions, other than in Aceh Province (Table 13). Ballesteros-Perez (2018) stated that the variable M1 was unintentionally delays realized by the contractors, and other stakeholders, which as an excusable delay to some extent for a range of time allowed.

Table 13

Results of RII for Natural Disaster Factors

Variable Code	Period I		Period II		Period III	
	RII	Assessment	RII	Assessment	RII	Assessment
M1	0.408	High	0.298	Medium	0.287	Medium
M2	0.136	Low	0.123	Low	0.101	Low
M3	0.219	Medium	0.081	Low	0.067	Low
M4	0.090	Low	0.075	Low	0.077	Low
M5	0.103	Low	0.102	Low	0.111	Low
M6	0.104	Low	0.087	Low	0.087	Low
M7	0.111	Low	0.091	Low	0.094	Low
M8	0.164	Medium	0.252	Medium	0.276	Medium
M9	0.099	Low	0.219	Medium	0.095	Low

Monetary Risk Assessment

Among monetary risk factors, the variable with the highest RII in Period I was N5 (economic crisis), in Period II was N4 (currency fluctuation), and in Period III was N3 (long-term valuation changes). Based on these conditions, it can be seen that these variables play an essential role in the delay of construction projects during all three study phases, as shown in Table 14.

Table 14
Results of RII for Monetary Factors

Variable Code	Period I		Period II		Period III	
	RII	Assessment	RII	Assessment	RII	Assessment
N1	0.115	Low	0.089	Low	0.081	Low
N2	0.132	Low	0.128	Low	0.140	Low
N3	0.149	Low	0.120	Low	0.136	Low
N4	0.171	Medium	0.136	Low	0.132	Low
N5	0.252	Medium	0.123	Low	0.134	Low

CONCLUSIONS

Risk schedule is influenced by variations in the frequency and severity described in the important scale. Risk variables with high scale can be considered as risk variables to be used as input in accessing risk schedule projects. This study has identified risk variables K4 (cultural conditions and local customs near project site), K5 (social issues/surrounding environment), L4 (increased fuel prices), M1 (uncertain weather conditions) as risk variables that hinder the achievement of project schedules in Period I. Variables L4 also need to be considered in Period II. The emergence of L4 variables in Period I was due to temporary security conditions, while in Period II due to the conditions of supply and demand.

Risk variables can appear at different periods, both in consecutive periods and over specified periods. The characteristics of the period influence the emergence of risk variables in these periods. Period I has security characteristics that are different from Period II which have project intensity characteristics as well as Period III which do not have the two previous characteristics, or we call them the “normal period”. However, the same risk variables can appear in different periods that are influenced by different causes.

The results of this study are intended to contribute to the application of risk to practitioners and governments. This research also provides ideas that can be used as a basis for evaluating the risks trend in further research for the other researchers.

ACKNOWLEDGEMENTS

The authors thank all members of the research team and the data provider for their support and helpful cooperation. We thank Bryan Schmidt from Edanz Group (www.edanzediting.com/ac) for editing this manuscript, and especially to the reviewers that have given a valuable suggestion, correction, and input for improvement of this paper for eligibility of publication.

REFERENCES

Abeku, D. M., Ogunbode, E. B., Salihu, C., Maxwell, S. S., & Kure, M. A. (2016). Projects management and the effect of rework on construction works: a case of selected projects in Abuja Metropolis, Nigeria. *International Journal of Finance and Management in Practice, 4*(1), 329-349.

Akintoye, A. S., & MacLeod, M. J. (1997). Risk analysis and management in construction. *International Journal of Project Management, 15*(1), 31-38.

Ariola, M. M. (2006). *Principles and methods of research*. Manila, Philippines: Rex Book Store.

Aven, T., & Renn, O. (2009). On risk defined as an event where the outcome is uncertain. *Journal of Risk Research, 12*, 1-11.

Ballesteros-Pérez, P., Smith, S. T., Papworth, J. G. L., & Cooke, P. (2018). Incorporating the effect of weather in construction scheduling and management with sine wave curves: application in the United Kingdom. *Construction Management and Economics, 36*(12), 666-682.

Banaitiene, N., & Banaitis, A. (2012). Risk management in construction projects. In N. Banaitiene (Ed.), *Risk management-currents issues and challenges* (pp. 429-448). Rijeka, Croatia: IntechOpen.

Bland, J. M., & Altman, D. G. (1997). Statistics notes: Cronbach's alpha. *British Medical Journal, 314*(7080), 572-572.

Christoplos, I., Mitchell, J., & Liljelund, A. (2001). Re-framing Risk: The Changing Context of Disaster Mitigation and Preparedness. *Disasters, 25*(3), 185-198.

Clayton, M. (2011). *Risk happens: Managing risk and avoiding failure in business projects*. London, UK: Marshall Cavendish International Asia Pte Ltd.

Construction Services Development Board. (2016). *Status of business entity registration process*. Retrieved February 25, 2016, from www.lpjk.net

Eduardus, I. S., & Hamsa, M. (2013). Service quality assesment in pt. indokemika jayatama using indserv scale and importance-performance analysis. *The Indonesian Journal of Business Administration, 2*(5), 543-553.

Fachrurrazi, Husin, S., & Mahmuddin. (2018). Project risk patterns: a comparison across three periods. *International Journal on Advanced Science, Engineering, Information Technology, 8*(5), 1997-2004.

Farrell, L. (2003). Principal-agency risk in project finance. *International Journal of Project Management, 21*(8), 547-561.

Fidan, G., Dikmen, I., Tanyer, A. M., & Birgonul, M. T. (2011). Ontology for relating risk and vulnerability to cost overrun in international projects. *Journal of Computing in Civil Engineering, 25*(4), 302-315.

Gray, C. F., & dan Larson, E. W. (2000). *Project management: The managerial process*. Singapore: McGraw-Hill.

Husin, S., Abdullah, A., Riza, M., & Afifuddin, M. (2017). Construction cost impact related to manpower, material, and equipment factors in contractor firms perspective. *AIP Conference Proceedings, 1903*(1), 1-7.

Husin, S., Abdullah, A., Riza, M., & Afifuddin, M. (2018). Risk assessment of resources factor in affecting project time. *Advances in Civil Engineering, 2018*, 1-9.

Kerzner, H. (2009). *Project management* (10th Ed.). New York, NY: John Willey and Sons.

Khodeir, L. M., & Mohamed, A. H. M. (2015). Identifying the latest risk probabilities affecting construction projects in Egypt According to political and economic variables From January 2011 to January 2013. *Housing and Building National Research Center, 11*, 129-135.

Latham, S. F., & Braun, M. (2008). Managerial Risk, Innovation, and Organizational Decline. *Journal of Management, 35*(2), 258-281.

Majid, A. M., & McCaffer, R. (1997). Assessment of work performance of maintenance contractors in Saudi Arabia. *Journal of Management in Engineering, 13*(5), 91.

Moazzami, M., Dehghan, R., & Ruwanpura, J. Y. (2011). Contractual risks in fast-track projects. *Procedia Engineering, 14*, 2552-2557.

Moe, T. L., & Pathranarakul, P. (2006). An integrated approach to natural disaster management. *Disaster Prevention and Management: An International Journal, 15*(3), 396-413.

Mubarak, Husin, S., & Oktaviati, M. (2017). External risk factors affecting construction costs. *AIP Conference Proceedings, 1903*(1), 1-9.

Pheng, L. S., Junying, L., & He, S. (2008). External risk management practices of chinese construction firms in Singapore. *KSCE Journal of Civil Engineering, 13*(2), 85-95.

Smith, M. J. D. (2015). *Statistical analysis handbook – A web-based statistics resource*. Winchelsea, UK: The Winchelsea Press.

Ward, S., & Chapman, C. (2003). Transforming project risk management into project uncertainty management. *International Journal of Project Management, 21*(2), 97-105.

Zeccola, P. (2011). Dividing disasters in Aceh, Indonesia: separatist conflict and tsunami, human rights and humanitarianism. *Disasters, 35*(2), 308-328.

Zhang, H. (2007). A redefinition of the project risk process: Using vulnerability to open up the event-consequence link. *International Journal of Project Management, 25*(7), 694-701.

Zhao, X., Wang, B. G., & Wang, W. (2013). Construction project risk management in Singapore: Resources, effectiveness, impact, and understanding. *KSCE Journal of Civil Engineering, 18*(1), 27-36.

Pertanika J. Sci. & Technol. 27 (4): 2049 - 2064 (2019)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Detection of Nitrobenzene using a coated Quartz Crystal Microbalance with a Parametric Modeling Approach

Dipali Ramdasi[1]* and Rohini Mudhalwadkar[2]

[1]*Department of Instrumentation and Control, Cummins College of Engineering for Women, Karvenagar, Pune, 411052, India*
[2]*Department of Instrumentation and Control, College of Engineering, Pune, 411005, India*

ABSTRACT

Nitrobenzene (NB) is a nitroaromatic compound possessing explosive properties. Timely detection of this compound will prevent emotional and financial losses to mankind. As direct sensors are not available, a parametric modeling approach is developed using the coated Quartz Crystal Microbalance (QCM). These sensors are mass sensitive and can measure changes in mass in few micrograms. When the sensing electrode is deposited with a chemically sensitive coating, adsorption of gas molecules occurs at the surface when exposed to a gas. This results in increased deposited mass and decrease in resonant frequency. This increase in mass is proportional to the gas concentration. A parametric model is derived for coated QCM using the Diffusion reaction model and kinetic analysis to estimate the gas concentration. When the concentration is beyond permissible limits, precautionary measures can be taken to prevent heavy losses. The parameters of sensing layer density and thickness, resonant frequency of quartz, exposition time were used to develop the model along with standard crystal parameters. The developed parametric model was validated experimentally using an AT cut 10MHz crystal, deposited with a sensitive coating of Polyaniline, Tungsten Oxide, Palladium and Dimethyl Sulfoxide. The GAMRY Electrochemical Quartz Crystal Microbalance (EQCM) was used for acquiring and analyzing the response of the QCM for NB sensing. The model was validated analytically with the Extended Butterworth Van Dyke (EBVD) model, rearranged for gas concentration determination. The model performed with an accuracy of 74% for concentrations higher than 300 ppb.

Keywords: Adsorption, chemically sensitive layer, diffusion, kinetic analysis, nitroaromatic explosive, nitrobenzene, parametric model, resonant frequency

ARTICLE INFO

Article history:
Received: 04 September 2018
Accepted: 23 April 2019
Published: 21 October 2019

E-mail addresses:
dipali.ramdasi@cumminscollege.in (Dipali Ramdasi)
rpm.instru@coep.ac.in (Rohini Mudhalwadkar)
* Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

INTRODUCTION

For years, countries like Iraq, Syria, India, Afghanistan and many more have experienced terrorist activities, leading to the death of hundreds of soldiers and innocent people every year. These countries bear a huge monetary loss for dealing with such incidents. According to a recently submitted report (Institute for Economics & Peace, 2018), India ranks 7th in the list of countries facing maximum terrorist attacks. The use of nitroaromatic explosives for destructive intentions is observed over the recent years. Thus, there is a need for timely detection of these explosives to avoid mishaps. Nitrobenzene (NB) (PubChem CID 7416) is an organic compound with the chemical formula $C_6H_5NO_2$. Mixtures with NB are found to be highly explosive with high sensitivity and detonation velocity as mentioned by Meyer et al. (2016), Giannoukos et al. (2016), Qingsong et al. (2007) and Woodfin (2006). The use of NB as an element of explosive mixtures was reported 50 years ago by Mason et al. (1965). It is a water insoluble pale-yellow oil with almond like odour. This paper deals with detection of NB as a potential explosive using a parametric modeling approach. As direct sensors for NB are not available, this method will result in the refinement of physical measuring techniques. For this, the Quartz Crystal Microbalance coated with a layer sensitive towards NB was used. The coated crystal was then subjected to varying concentrations of NB and the response of the sensor in terms of deviation in resonant frequency was observed. The concentration of NB was then estimated from the developed model substituting the observed sensor response. The basic properties of QCM, thin film sensors and the method adopted for developing the model are further discussed.

MATERIALS AND METHODS

Quartz Crystal Resonator

When the Quartz Crystal Resonator (QCR) is deposited with material of a very low mass of a few micrograms, a change in the resonating frequency is observed as stated by Fraden (2015) and Vives (2008). The deposited mass is proportional to change in resonant frequency and can be calculated using the Sauerbrey's equation as proposed in an earlier article (Sauerbrey, 1959).

$$\Delta f = \frac{-2f_0^2 \Delta m}{A\sqrt{\rho_q G_q}} \qquad [1]$$

This property of QCR can be used for detecting a Volatile Organic Compound (VOC) if the change in mass is due to a VOC. For this, the QCM is deposited with a polymer based chemically sensitive coating as presented by Reglero et al. (2018), Temel and Tabakci (2016), Bai and Shi (2007), Seena et al. (2011) and Menon et al. (1998) and is shown in Figure 1.

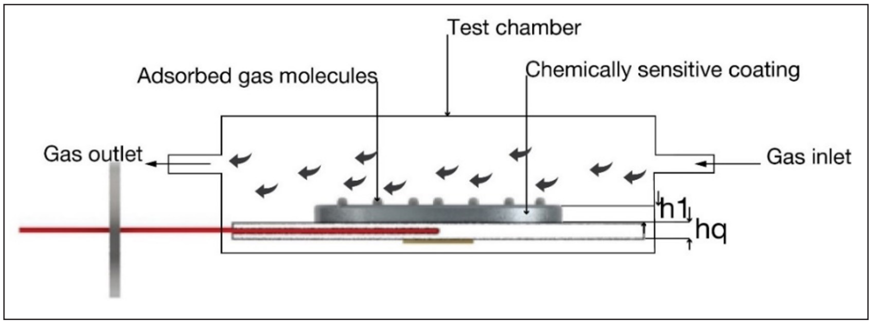


Figure 1. Adsorbed gas molecules on the chemically sensitive coating of QCM

Table 1
Properties of AT cut Quartz crystal

Quartz Parameter	Value	Description
ρ_q	2651 Kg/m^3	density of quartz crystal
η_q	9.27 *10^{-3} Pa. s	viscosity of quartz crystal
G_q	2.93*10^9 Pa	shear modulus of AT cut quartz crystal
f_0	10 MHz	resonant frequency of crystal
ε_{22}	3.982*10^{-11} A^{2}s^4/kgm^3	permittivity of quartz
c_{66}	9.2947 * 10^{10} N/m^2	component of material property tensor for mechanical stiffness
Z_{cq}	8852147 Ω	Characteristic impedance of quartz crystal

This sensor structure is similar to thin film sensors, with the QCR as the substrate. Thus, laws of diffusion as applied to thin film sensors are applied to the coated QCM along with kinetic analysis of the QCM. Coated QCM has been used in liquids as presented by Arnau et al. (2001), Rabe et al. (2003) and Lucklum et al. (1999). Its electrical equivalent circuit is represented by the Extended Butterworth Van Dyke model (EBVD) as proposed by Arnau et al. (2001) considering shear parameters as used by Johannsmann et al. (1992). The elements of the EBVD model are considered frequency independent as mentioned by Martin et al. (1991). Limited literature is published on use of QCM for nitroaromatic gas concentration estimation.

An AT cut quartz crystal in thickness shear mode was coated and placed in a test chamber of known volume. Adsorption of gas molecules occurs on the coated surface proportional to concentration, which increases the deposited mass. Kinetic analysis of QCM was combined with Diffusion Reaction Model (DRM) to develop the parametric model. Some values of the QCM considered for modeling are illustrated in Table 1 as presented by Arnau et al. (2001).

The equivalent circuit of the quartz crystal loaded with a finite viscoelastic layer contacting a semi-infinite viscoelastic medium is illustrated in Figure 2 and known as the

Extended Butterworth Van Dyke (EBVD) Model. Changes in the surface mass density of coating produce changes in the electrical impedance of the sensor.

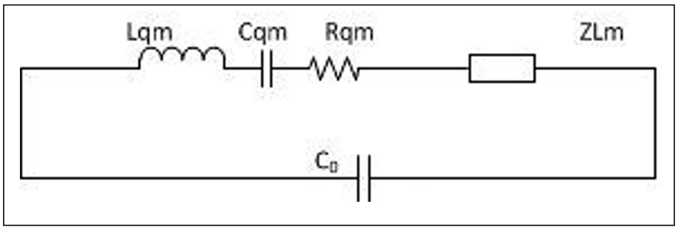


Figure 2. Circuit representation of the EBVD model

The complex electrical input impedance Z between terminals of loaded QCR can be expressed as:

$$Z = \frac{1}{j\omega C_0}\left(1 - \frac{K_q^2}{\alpha_q} . \frac{2\, tan\frac{\alpha_q}{2} - j\frac{Z_l}{Z_q}}{1 - j\frac{Z_L}{Z_q}cot\alpha_q}\right) \qquad [2]$$

Where α_q = acoustic phase shift in quartz crystal, K_q^2=complex electromechanical coupling factor = $K_q^2 = \frac{e_q}{\varepsilon_q c_q}$. The characteristic quartz impedance $Z_q = \sqrt{c_q \rho_q}$ where and $c_q = \overline{c_{66}} = c_{66} + \frac{e_{26}^2}{\varepsilon_{22}}$ is the piezoelectric stress coefficient. The values of circuit elements of the EBVD model are given by $R_m^q = \frac{(N\pi)^2 n_q}{8K_q^2 C_0 c_q}$, $C_m^q = \frac{8K_q^2 C_0}{(N\pi)^2}$, $L_m^q = \frac{1}{\omega_s^2 C_0}$ and $C_0 = \frac{\varepsilon_{22} A}{d_q}$ where ω_s is the series resonating frequency, d_q is the thickness of quartz disc and A is area of electrodes. The surface mechanical load impedance for a single coating which is equivalent to the motional load impedance Z_L exposed to a gas can be stated with equation [3] and proposed by Lucklum et al. (2006).

$$Z_L = j\sqrt{\rho_1 G_1 tan\left(h_1 \omega \sqrt{\frac{\rho_1}{G_1}}\right)} \qquad [3]$$

where h_1 = thickness as shown in Figure 1 and ρ_1 is density of deposited layer, G_1 is complex elastic modulus of sensing layer and expressed as $G_1 = G_1{}' + jG_1{}''$, $G_1{}'$ where =Energy storage and $G_1{}''$ = Energy dissipation in layer. Loss angle δ_1 is expressed as $\tan \delta_1 = \frac{G_1{}''}{G_1{}'}$. To analyze the response of QCM, the parameter change in resonant frequency needs to be measured and is expressed as equation [4].

$$\Delta f_s \approx \left(\frac{f_s}{N\pi} \sqrt{\frac{G_2''\rho_2}{2\rho_q \bar{c}_{66}}} \left(\frac{\sqrt{1+\tan^2\delta_2}-1}{\tan\delta_2} \right)^{\frac{1}{2}} + \frac{2f_s^2\rho_s}{N\sqrt{\rho_q \bar{c}_{66}}} \right) \quad [4]$$

G_2'' is the shear contribution and ρ_2 is the density of gas under test. δ_2 is the loss angle between sensing layer and medium under test. As the layer 2 is considered gas in the developed model, the EBVD model cannot be used for gas concentration estimation as the phenomenon of adsorption and desorption is not included. Also, as the EBVD model is a standard model, the performance of a newly developed model can be evaluated by comparing its results with the EBVD model. However, if the newly developed model is developed for a target gas, the EBVD model needs to be rearranged.

Thin Film Sensors

Adsorption and desorption of gas molecules due to diffusion occurs on the chemically sensitive coating, resulting in change in electrical properties of the sensor. The sensor was modeled using DRM and configuration of the sensor is as shown in Figure 3, where ‘x’ is the thickness of the film. ‘x’=L indicates that the thickness of the film is L.

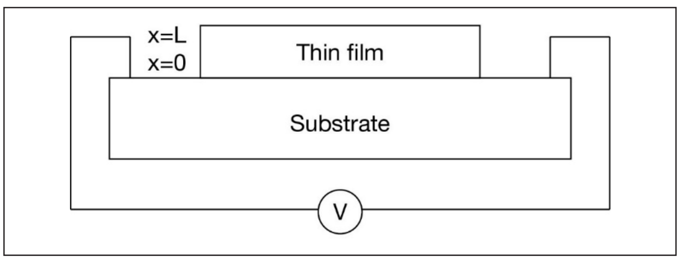


Figure 3. Thin film sensors

The gas adsorption to the surface of thin film sensors is based on Langmuir Isotherm and is expressed as $A + S \overset{k_f}{\leftrightarrow} AS$ where A, S and AS are the detected gas, the polymer film free surface area, and the occupied surfaces of the sensing film, respectively. k_f and k_b are the rate of forward and backward adsorption kinetics. Taking *x* as the depth of adsorption in the film, t as the time and $\theta(\chi, \tau)$ as the fraction of occupied surfaces on sensing film, the diffusion reaction equation for species of gas ‘a’ was presented by Galic et al. (2006) and Gardner et al. (1995).

$$D\frac{\partial^2 a}{\partial x^2} - k_f\, a(1-\theta)N + k_b\theta N = \frac{\partial a}{\partial t} \quad [5]$$

In equation [5], concentration of diffusing gas ‘a’ is a function of *x* and time. *D* is the diffusion rate and *N* is the density of film. L is the thickness of sensing film as shown in

Figure 3 and a_∞ is the maximum concentration of gas. Using the boundary conditions as t=0; θ = 0 and $0 \le x \le L$, then at t > 0; $\frac{da}{dx}$ =0 at x = 0 and a = a_∞ at x = L. Converting it to dimensionless form, an approximate analytical solution for linear diffusion in dimensionless form is expressed in equation [6].

$$\eta \frac{\partial\theta}{\partial\tau} = \kappa\lambda\gamma - \kappa\theta \qquad [6]$$

Normalized gas concentration $\gamma = \frac{a}{a_\infty}$; τ is the dimensionless time parameter given by $\tau = \frac{t}{t_{cycle}}$ where t is the time, $t = t_{cycle}$ is the time for which sensor is exposed to gas. $\eta = KN$ where $K = \frac{k_f}{k_b}$, N= density of film and $\lambda = Ka_\infty$. The dimensionless depth of adsorption $\chi = \frac{x}{L}$; and $\kappa = \frac{k_f NL^2}{D}$. The gas concentration profile $\gamma(\chi,\tau)$and occupied surfaces $\theta(\chi,\tau)$ are expressed in equation [7] and [8] as proposed by Gardner et al. (1995) and Bartlett and Gardner (1996).

$$\gamma(\chi,\tau) = 1 - \frac{2}{\pi}\sum_{n=0}^{\infty} \frac{\cos[(\eta + 1/2)\pi\chi]\exp\left[-(\eta + \frac{1}{2})^2\pi^2\tau\right]}{(-1)^\eta(\eta + 1/2)} \qquad [7]$$

$$\theta(\chi,\tau) \approx \lambda\gamma \qquad [8]$$

This study summarizes that the surfaces occupied with gas molecules are dependent on the γ, k_f, k_b, a_∞ and ρ_1.

Model Development

The kinetic analysis is a study depending on the relative and absolute concentration of analyte with time. When the sensing layer is subjected to the gas under test, the adsorption progresses, resulting in increase of deposited mass and reduction in resonant frequency. The phenomenon is explained with the equation [9].

$$\theta(t) \propto \Delta m(t) \propto -\Delta f(t) \qquad [9]$$

The occupied surface on the sensing film $\theta(t)$ depends on the following parameters; viz. gas concentration, forward and backward reaction rate, maximum concentration of the gas under test in the test environment, density of sensing layer and the time for which the sensor is subjected to gas. The $\Delta m(t)$ is proportional to change in frequency, which in turn is expressed in equation [10].

$$\Delta m(t) \propto \frac{Im(Z_L)(t)}{\pi Z_{cq}} \qquad [10]$$

Where Z_L is the acoustic load impedance given by equation [3] and $Z_{cq}$is the characteristic quartz impedance and calculated using the R_m^q, C_m^q, L_m^q and C_0. Applying the boundary conditions of depth of adsorption x as h_1, which is the thickness of sensing film and time for which the sensor is subjected $t = t_{cycle}$, the equation [9] is solved further. At t= 0; θ(t) = 0; and 0 ≤ depth of adsorption x ≤ h_1, then at t > 0;

$$\frac{da}{dx} = 0;\text{ at } x = 0 \text{ and } t = 0;\ a = a_\infty \text{ at } x = h_1 \text{ and } t = t_{cycle} \qquad [11]$$

a is the concentration of the gas and a_∞ is the maximum gas concentration. From the equation of occupied surface used in the DRM and the kinetic analysis of the QCM, the equation [9] is solved with above stated boundaries and a parametric model of DRM and kinetic analysis of QCM is derived for estimating the gas concentration.

$$a(t) = \frac{C.a_\infty f_0}{\lambda \pi Z_{cq}} \sum_{t=0}^{t=t_{cycle}} \frac{-Im(Z_L)(t)}{\Delta f(t)} \qquad [12]$$

In the above model, $\lambda = Ka_\infty,\ K = \frac{k_f}{k_b}$, Z_L is calculated as expressed in equation [7], where h_1 is considered as 16.61 μm. G_1 is complex elastic modulus of sensing layer and C is a fitting constant. $\Delta f(t) = f_L(t) - f_0$, where f_L is the resonant frequency of the crystal when subjected to gas at $t=0$ to $t= t_{cycle}$ and f_0 is the resonant frequency of unperturbed quartz crystal. $G_1' = 10^9$ Pa and $G_1'' = 10^2$ Pa and considered same for the above model.

The following assumptions are done while developing the parametric model

- Surface load mechanical impedance can be represented with the density and thickness of the sensing layer along with shear parameters
- The layer deposited is very thin and is free of any prior adsorption
- Homogenous adsorption of gas molecules throughout the surface
- Adsorption and desorption cycle time are approximately same with desorption starting by step change in gas concentration to zero

Experimentation

A comprehensive study was done from the work of Cui et al. (2016), Palaniappan et al. (2008), Özmen et al. (2006), Yang and He (2016), Erbahar et al. (2011), Li et al. (2007) and Ramdasi and Mudhalwadkar (2018) for identifying the materials of sensing layer. As the performance of polymers with metal oxides for explosive detection was prominent, polyaniline (Sigma-Aldrich, India) based sensing layer with palladium metal powder (Ottokemi, India) for selectivity enhancement was used. The following procedure was adopted to develop the chemically sensitive layer. A micro balance was used to weigh the powders accurately. The weight of the powders depends on the pre-decided ratio of the mixture of powders to achieve a higher selectivity and sensitivity. The weight ratios decided for the sensor was a mixture of selective material powder (6%) and sensitive material powder (94%). The sensitive powder was a mixture of Polyaniline (65%) Tungsten Oxide (Research-Lab Fine Chem Industries, India) (35%). As the powders needed to mix properly and micro or nano granules to form, severe pulverization of the mixture was done. The pulverized powder was mixed with 2-3 drops of glycerin (Sigma-Aldrich, India) which acted as a binding agent. 1-2 drops of Dimethyl Sulfoxide (Sigma-Aldrich, India) were also added, which acted as a chemical agent. Small part of the mixture was carefully placed on the crystal with a micropipette and the crystal was then sintered at 80°C in an oven for 40 minutes. It was observed that if the mixture deposited on the crystal becomes excess, the crystal ceased to resonate. The density of the layer was calculated using formula for density of mixture and accounts to 0.804 g/cm^3.

Figure 4 shows the uncoated 10 MHz Au coated quartz crystal (971-00006, GAMRY Instruments, USA) used for experimentation while Figure 5 shows the crystal after drop coating. Figure 6 shows the coated crystal after sintering. The distance between the two terminals is 9 mm, which is approximately equal to the quartz crystal resonators.

The a_∞ was calculated using the upper explosive limit of NB and density of NB was 1.2 g/cm^3. The volume of test chamber was 560 ml. The forward and backward reaction rate was 0.65034/s and 0.45563/s respectively as suggested in the literature. The thickness of the crystal was 0.00016 m and electrode area was 0.293 cm^2. The standard crystal parameters proposed by Arnau et al. (2001) were considered and are detailed in Table 1. The experimental setup included an air tight chamber with inlet and outlet connections

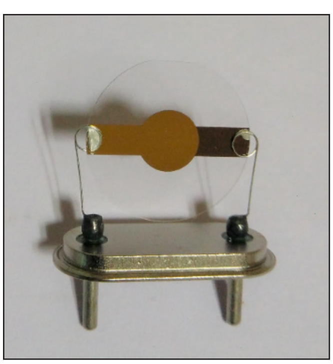


Figure 4. Uncoated AT cut 10 MHz quartz crystal

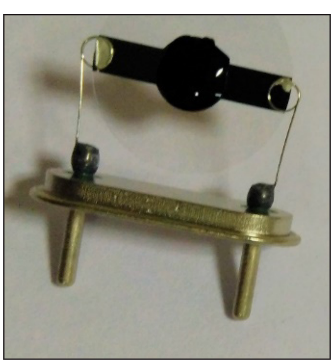


Figure 5. Crystal after drop coating

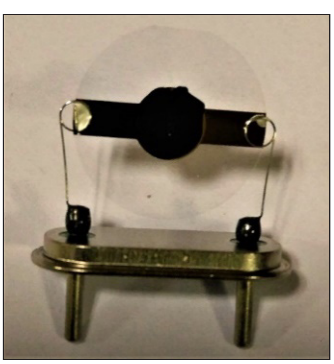


Figure 6. Coated crystal used for experimentation after sintering

for gas under test. The chamber was fitted with a cuvette which held the coated QCR connected to a computer via the Electrochemical Quartz Crystal Microbalance (EQCM) (eQCM10M, GAMRY Instruments, USA). The resonant frequency of 10 MHz changes to 9.995 MHz after deposition of the chemically sensitive layer. The concentration of gas was calculated for a variation in resonant frequency from 9.995 MHz to 9.992 MHz for a cycle time of 600 seconds.

Model Validation by Experimental Method

The developed model was validated experimentally by subjecting the coated QCM to a varying concentration of NB from 30 ppb to 1000 ppb. The response of the coated QCM was acquired and analyzed using the GAMRY EQCM resonator software in the Echem analyst environment. For this, the coated QCM was placed in an air-tight chamber of 560 mL volume, in a cuvette as illustrated in Figure 7 and was exposed to known concentration of NB via a pump (20K series, Boxer Pump, UK) with speed set to 400 mL/ min. If the speed of the pump was lower, the sensing layer required a higher time to saturate, thus increasing the response time of the sensor. As the sample gas came in and left the gas chamber, the pressure in the gas chamber was assumed to be at atmospheric value. The complete setup of experimentation is presented in Figure 8.

To generate the desired concentration, the calculated amount of volatile compound was introduced in liquid phase in an air tight container. The compound was allowed to mix with the air in the container. The following procedure was adopted to calculate the volume to be injected in liquid phase.

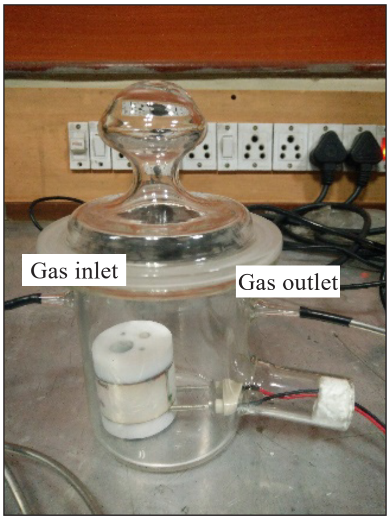


Figure 7. Quartz Crystal placed in cuvette and exposed to NB

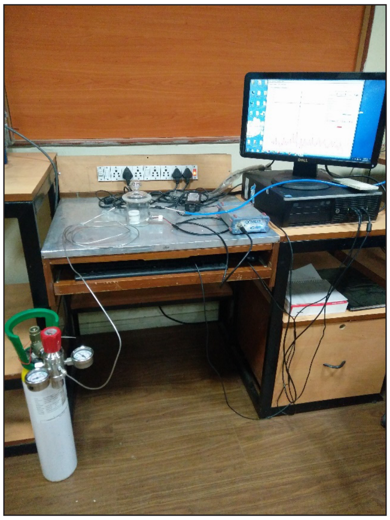


Figure 8. Experimental setup for coated Quartz Crystal Microbalance

a. Calculate the density ∂ in g/litre using the following formula and values

$$\partial = \frac{P.MW}{R.T} \qquad [13]$$

where P is the standard atmospheric pressure, 1 atm;
MW is the molecular weight of compound (NB=123.06 g/mol);
R is the Universal gas constant in L.atm/mol.K = 0.0821
T is the temperature in degree Kelvin

b. Using the formula, calculate the volume to be injected in liquid phase

$$V_{liq} = \frac{V_{gas}.\partial}{\rho} \qquad [14]$$

where V_{gas} is the required volume of compound in gaseous state;
ρ is the density in g/cm^3 (NB= 1.2)

The change in resonant frequency and the corresponding addition in mass was noted using the e-QCM software and its corresponding concentration in Parts Per Billion (ppb) was calculated using standard formula.

Model Validation by Analytical Method

The results of the derived model were compared with the EBVD model. The equation for change in resonant frequency was modified for measuring gas concentration considering volume of test chamber and density of NB and stated in equation [15].

$$m(t) = \frac{2\rho_q \bar{c}_{66}}{G_2''}\left(\frac{N\pi V^2}{f_0}\right)^2 \left(\frac{\tan\delta_2}{\sqrt{1+\tan^2\delta_2}-1}\right)\left(\Delta f(t) + \frac{2f_0^2\rho_1 h_1}{N\rho_q\bar{c}_{66}}\right)^2 \qquad [15]$$

In the equation, δ_2 is the loss angle between the sensing layer and gas layer, considered as 1.57 and $G_2'' = 1.3 * 10^3$ Pa. The deviation in resonant frequency due to adsorption of gas molecules on the chemically sensitive coating is considered as 0 to 3000 Hz.

RESULTS AND DISCUSSION

A parametric model for a coated QCM is derived using DRM and kinetic analysis of QCM. The results are compared with the rearranged equation of EBVD model.

As the DRM is governed by adsorption, a study of the diffusion of gas molecules due to adsorption was done. Figure 9 shows the normalized concentration during adsorption for chemically sensitive coated sensor. Increase in normalized concentration was observed when the depth of adsorption in the coating increased for varying time cycles.

Figure 10 shows that as the time for which the sensor was exposed to the gas under test progresses, more surface of the sensing layer was occupied due to adsorption of gas, till the surface neared saturation. Equation [8] explains this phenomenon, which is further used for kinetic analysis of the coated QCM. This phenomenon was also observed during the experimental validation. When the coated QCM was exposed to gas, the sensing layer surface experiencing adsorption increased, which in turn increased the mass deposited on the layer. This ultimately resulted into a reduced resonant frequency of the coated quartz crystal. The developed parametric model was validated experimentally as well as with the rearranged EBVD model. Figure 11 illustrates the response of the developed parametric model for detection of NB along with experimental and analytical validation. The developed model and the experimentation indicate that, for a 10 MHz crystal coated with a sensing layer for NB, the deviation in the resonant frequency is a measure of concentration of the gas. Thus, the X axis of Figure 11 represents the deviation in the frequency for 10 MHz crystal and the Y axis indicates the corresponding gas concentration in ppb. The graph is a representation of the measurement system results for lower concentration range of NB, from 30 ppb to 1000 ppb. The figure shows the actual concentration of NB in parts per billion for the respective change in resonant frequency for the developed model, the modified EBVD model and experimentation. The graph exhibits a nonlinear relation with the deviation in resonant frequency due to nonlinear nature of the phenomenon of adsorption. For a maximum deviation in frequency of 3000 Hz, the concentration of NB at $t= t_{cycle}$ is 978 ppb for EBVD model while it is 976 ppb for the derived parametric model for an actual concentration of 1000 ppb. The developed model thus performs with an accuracy of 74% for NB concentrations above 300 ppb. As the EBVD is an electronic model, it lacks the effect of adsorption, due to which the nature of the graphs of both models is different.

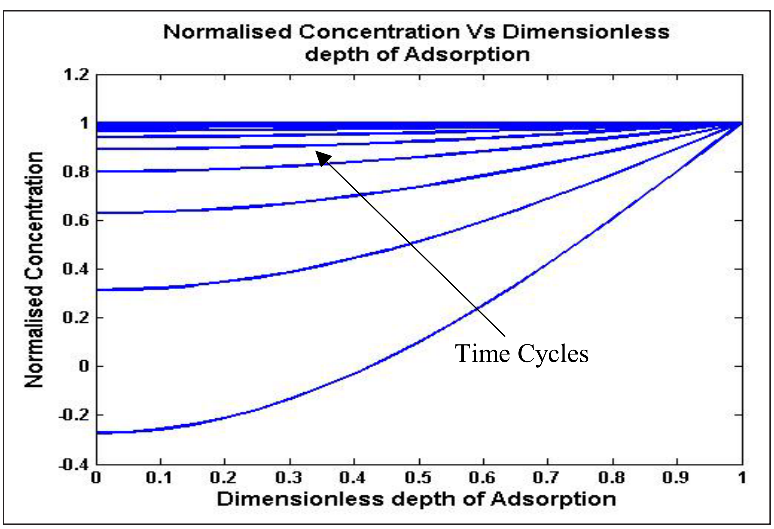


Figure 9. Normalised concentration vs Depth of adsorption

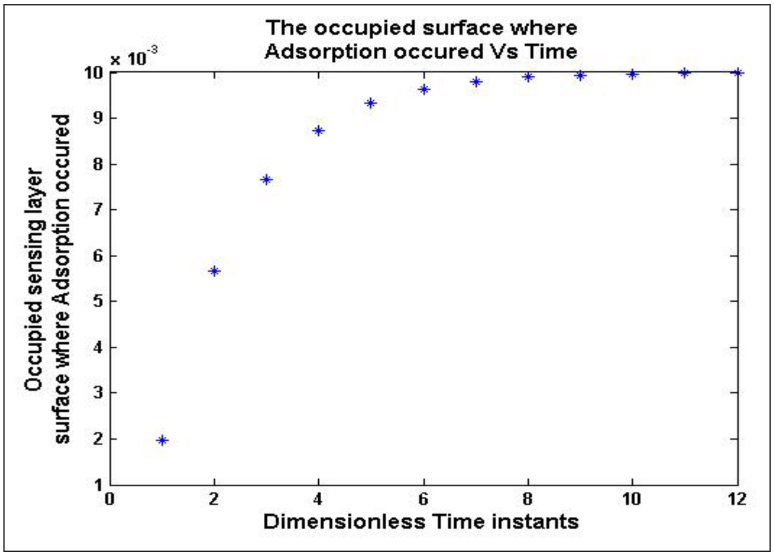


Figure 10. Occupied Surface vs Time for which sensor was exposed to gas

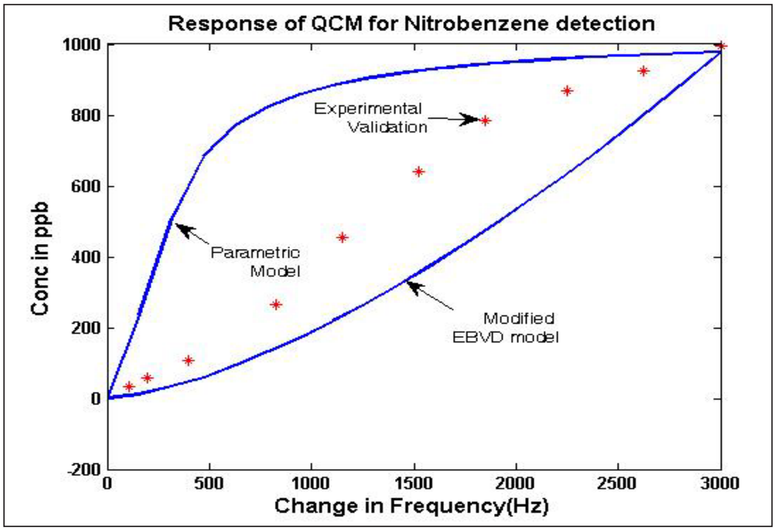


Figure 11. Concentration estimation of nitrobenzene using parametric model and validation results with experimental and analytical method

Figure 12 shows response of the coated QCM when subjected to a concentration of 800 ppb nitrobenzene, as observed on the EQCM software. The EQCM software is a powerful tool to analyze the behavior of the quartz crystal, when subjected to deposition of mass on the electrode surface. The X axis represents the time instants, when the coated QCM is subjected to gas. The graph has two Y axis, one axis represents the deviation in the resonant frequency of 10 MHz crystal, indicated by the 'dF series'. The second Y axis indicates the mass deposited on the quartz crystal in µg, indicated by deltaM. The observed mass is 2.75

µg in the volume of 560 ml. The concentration in ppm is calculated by first calculating the mass in one-liter volume and then using equation [16]. Standard atmospheric temperature and pressure are used for calculations.

$$Conc_{ppm} = \frac{mass\ per\ liter * Atmospheric\ Pressure}{Molecular\ wt\ of\ NB\ in\ g/mol} \qquad [16]$$

As the process of adsorption and desorption are nonlinear, a fitting constant is incorporated in the model. The sensing layer is chemically sensitive to NB and developed from Polyaniline, Tungsten Oxide, Palladium and Dimethyl Sulfoxide.

The explosives exhibit nitroaroma, however reliable sensors are not available to detect nitroaroma in low level of detection, typically ppb. This model uses the advantage of mass sensitivity of the quartz crystals to detect small changes in the mass deposited on the quartz surface. The response time of the sensor is observed as 6.5 seconds while the recovery time is 180 seconds. Higher response and recovery time is a characteristic of thin film sensors, which is observed in the developed sensor also, due to its functional relevance with thin film sensors. The change in deposited mass is an effect of adsorption of gas molecules on the sensitive layer surface. With detailed study and knowledge of sensitive coating materials for gas detection, exact coatings of chemically sensitive layers can be developed by different techniques. This model than can assist in estimation of gas concentration and can be further extended to decide the presence of explosives. The timely detection of explosives can prevent heavy losses to mankind and nature.

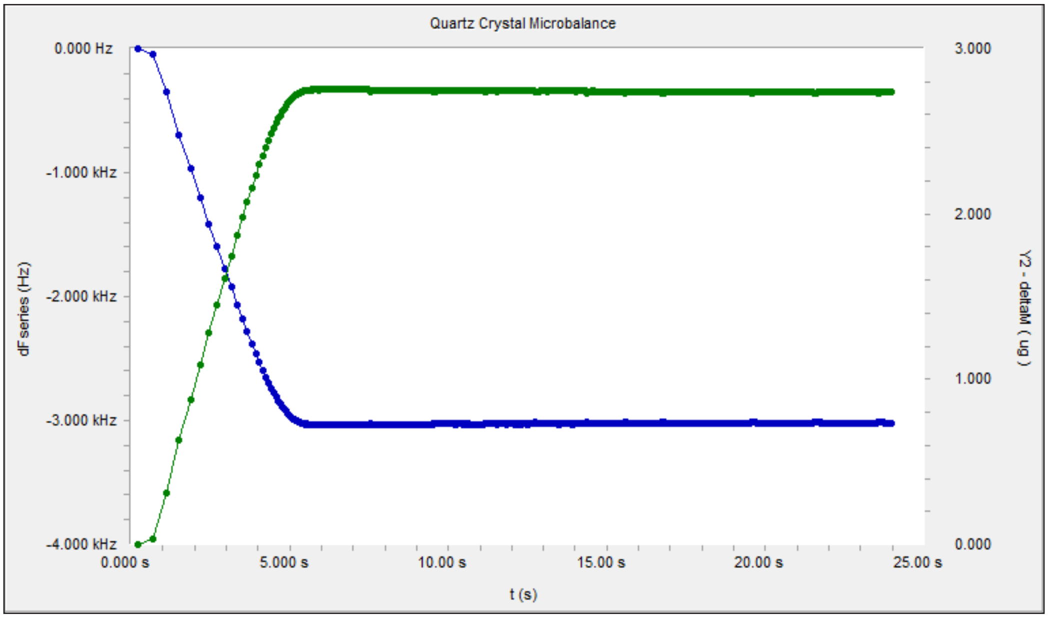


Figure 12. Response of coated QCM for 800 ppb nitrobenzene

CONCLUSION

A parametric model for detecting NB was developed for a 10 MHz AT cut quartz crystal deposited with a chemically sensitive coating. The parameters of cycle time, density and thickness of sensing layer were considered for modeling along with standard crystal parameters. The results of the developed model were compared with the rearranged equation of Extended Butterworth Van Dyke Model. The model was validated experimentally for concentration of NB ranging from 30 ppb to 1000 ppb. For concentrations lower than 300 ppb, the derived model exhibits maximum deviation, however, for concentrations above 300 ppb, the derived model estimated the concentration of NB with an accuracy of 74%. The mass sensitivity and advantage of chemical selectivity using chemically sensitive layer was combined in the derived model. This model can assist researchers in the selection of elements of the chemically sensitive coating for a target gas and leads one step closer to timely detection of explosives.

ACKNOWLEDGEMENT

This project is funded by Bhabha Atomic Research Centre, India under BRNS scheme, file number 2012/36/55/BRNS/March 2013

REFERENCES

Arnau, A., Jimenez, Y., & Sogorb, T. (2001). An extended Butterworth Van Dyke model for quartz crystal microbalance applications in viscoelastic fluid media. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, 48*(5), 1367-1382.

Bai, H., & Shi, G. (2007). Gas sensors based on conducting polymers. *Sensors*, *7*(3), 267-307.

Bartlett, P. N., & Gardener, J. W. (1996). Diffusion and binding of molecules to sites within homogeneous thin films. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences, 354*(1704), 35-57.

Cui, S., Yang, L., Wang, J., & Wang, X. (2016). Fabrication of a sensitive gas sensor based on PPy/TiO2 nanocomposites films by layer-by-layer self-assembly and its application in food storage. *Sensors and Actuators B: Chemical, 233*, 337-346.

Erbahar, D. D., Gürol, I., Ahsen, V., Öztürk, Z. Z., Musluoğlu, E., & Harbeck, M. (2011). Explosives detection in sea water with phthalocyanine quartz crystal microbalance sensors. *Sensor Letters, 9*(2), 745-748.

Fraden, J. (2015). *Handbook of modern sensors: physics, designs, and applications.* New York, NY: Springer Science & Business Media.

Galic, N., Vujovic, I., & Kuzmanic, I. (2006, June 7-10). Thin film gas sensor modeling and simulation. In *Proceedings ELMAR 2006* (pp. 239-242). Zadar, Croatia.

Gardner, J. W., Bartlett, P. N., & Pratt, K. F. E. (1995). Modelling of gas-sensitive conducting polymer devices. *IEE Proceedings-Circuits, Devices and Systems, 142*(5), 321-333.

Giannoukos, S., Brkić, B., Taylor, S., Marshall, A., & Verbeck, G. F. (2016). Chemical sniffing instrumentation for security applications. *Chemical reviews, 116*(14), 8146-8172.

Institute for Economics & Peace. (2018, November). *Global Terrorism Index 2018: Measuring the impact of terrorism*. Retrieved November 01, 2018 from http://visionofhumanity.org/app/uploads/2018/12/Global-Terrorism-Index-2018-1.pdf

Johannsmann, D., Mathauer, K., Wegner, G., & Knoll, W. (1992). Viscoelastic properties of thin films probed with a quartz-crystal resonator. *Physical Review B, 46*(12), 7808-7815.

Li, G., Zheng, J., Ma, X., Sun, Y., Fu, J., & Wu, G. (2007). Development of QCM trimethylamine sensor based on water soluble polyaniline. *Sensors, 7*(10), 2378-2388.

Lucklum, R., Behling, C., & Hauptmann, P. (1999). Role of mass accumulation and viscoelastic film properties for the response of acoustic-wave-based chemical sensors. *Analytical Chemistry, 71*(13), 2488-2496.

Lucklum, R., Doerner, S., Schneider, T., Schlatt-Masuth, B., Jacobs, T., & Hauptmann, P. (2006, June 4-7). Real time kinetic analysis with quartz crystal resonator sensors. In *2006 IEEE International Frequency Control Symposium and Exposition* (pp. 528-534). Miami, FL, USA.

Martin, S. J., Granstaff, V. E., & Frye, G. C. (1991). Characterization of a quartz crystal microbalance with simultaneous mass and liquid loading. *Analytical Chemistry, 63*(20), 2272-2281.

Mason, C. M., Van Dolah, R. W., & Ribovich, J. (1965). Detonability of the system nitrobenzene, nitric acid, and water. *Journal of Chemical and Engineering Data, 10*(2), 173-175.

Menon, A., Zhou, R., & Josse, F. (1998). Coated-quartz crystal resonator (QCR) sensors for on-line detection of organic contaminants in water. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, 45*(5), 1416-1426.

Meyer, R., Köhler, J., & Homburg, A. (2016). *Explosives*. Weinheim, Germany: John Wiley & Sons.

Özmen, A., Tekce, F., Ebeoğlu, M. A., Taşaltın, C., & Öztürk, Z. Z. (2006). Finding the composition of gas mixtures by a phthalocyanine-coated QCM sensor array and an artificial neural network. *Sensors and Actuators B: Chemical, 115*(1), 450-454.

Palaniappan, A., Moochhala, S., Tay, F. E., Su, X., & Phua, N. C. (2008). Phthalocyanine/silica hybrid films on QCM for enhanced nitric oxide sensing. *Sensors and Actuators B: Chemical*, *129*(1), 184-187.

Qingsong, J. W., Song, G., & Jinghua, S. (2007). Nitrobenzene and aniline caused fire and explosion: a case study. In *Institution of Chemical Engineers Symposium Series* (Vol. 153). Rugby, England: IChemE.

Rabe, J., Buttgenbach, S., Schroder, J., & Hauptmann, P. (2003). Monolithic miniaturized quartz microbalance array and its application to chemical sensor systems for liquids. *IEEE Sensors Journal, 3*(4), 361-368.

Ramdasi, D., & Mudhalwadkar, R. (2018). Thin film sensor materials for detection of Nitro-Aromatic explosives. *IOP Conference Series: Materials Science and Engineering, 323*(1), 1-6.

Reglero Ruiz, J., Sanjuán, A., Vallejos, S., García, F., & García, J. (2018). Smart polymersin micro and nano sensory devices. *Chemosensors, 6*(2), 12-31.

Sauerbrey, G. Z. (1959). The use of quarts oscillators for weighing thin layers and for microweighing. *Journal for Physics*, *155*, 206-222.

Seena, V., Fernandes, A., Pant, P., Mukherji, S., & Rao, V. R. (2011). Polymer nanocomposite nanomechanical cantilever sensors: material characterization, device development and application in explosive vapour detection. *Nanotechnology*, *22*(29), 1-11.

Temel, F., & Tabakci, M. (2016). Calix [4] arene coated QCM sensors for detection of VOC emissions: Methylene chloride sensing studies. *Talanta, 153*, 221-227.

Vives, A. A. (Ed.). (2008). *Piezoelectric transducers and applications.* Heidelberg, Germany: Springer-Verlag.

Woodfin, R. L. (Ed.). (2006). *Trace chemical sensing of explosives.* Hoboken, NJ: John Wiley & Sons.

Yang, M., & He, J. (2016). Graphene oxide as quartz crystal microbalance sensing layers for detection of formaldehyde. *Sensors and Actuators B: Chemical, 228*, 486-490.

Pertanika J. Sci. & Technol. 27 (4): 2065 - 2077 (2019)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Comparison of Overall Heat Transfer Coefficient between Shell and Tube and Spiral Coil Heat Exchangers

Meseret Nasir Reshid[1]*, Girma Tadesse Chala[2] and Wan Mansor Wan Muhamad[1]

[1]*Department of Mechanical Engineering, University Kuala Lumpur Malaysia France Institute Section 14, Jalan Teras Jernang, 43650 Bandar Baru Bangi, Selangor, Malaysia*

[2]*International College of Engineering and Management, P.O. Box 2511, C.P.O Seeb, P.C 111, Muscat, Oman*

ABSTRACT

Heat exchangers are used in many industries and power generation applications. The performance of heat exchangers depends on the operating parameters and the types of flow. The sudden pressure drop is one of the major problems encountered in heat exchanger, and this would significantly affect the efficiency and the overall heat transfer coefficient of the heat exchanger. Therefore, this study is aimed at investigating and analyzing the effects of operating parameters that cause pressure fluctuation and affect the overall heat transfer coefficient. Experimental study was carried out for two types of flows: co-current and counter concurrent flows. Comparisons of the overall heat transfer coefficients between shell and tube and spiral coil heat exchangers were made. It was observed that mass flow rate affected the overall heat transfer coefficient. Besides, the counter current flow was more efficient compared to the co-current flow with enhanced overall heat transfer coefficient. The maximum overall heat transfer coefficient for spiral coil heat exchanger counter flow was 2702.78 W/m^2.K, showing a higher heat transfer efficiency when compared to the shell and tube heat exchanger. Moreover, the spiral coil heat exchanger occupied less space as opposed to the shell and tube heat exchanger.

Keywords: Co-current, counter current, overall heat transfer coefficient, pressure drop, shell and tube heat exchanger, spiral coil heat exchanger

ARTICLE INFO

Article history:
Received: 28 August 2018
Accepted: 25 June 2019
Published: 21 October 2019

E-mail addresses:
meseret@unikl.edu.my; meseretreshid@gmail.com (Meseret Nasir Reshid)
girma_tade@yahoo.com (Girma Tadesse Chala)
drwmansor@unikl.edu.my (Wan Mansor Wan Muhamad)
*Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

INTRODUCTION

Heat exchangers have been used in different engineering applications since the beginning of civilization (Zohuri, 2017). Its wide variety of applications include heat recovery systems, food industries, natural gas processing, space heating, sewage treatment, power stations, chemical plants, petrochemical plants refrigeration and air conditioning (Mohanraj et al., 2015; Ma'arof et al., 2019). There have been many techniques used for heat transfer enhancement in industrial applications, among which active and passive techniques were the two major enhancement techniques (Seyed, 2013). The heat transfer enhancement enables the size of the heat exchanger to be small, and improves the performance of heat exchanger. Faramarz (2013) compared active and passive enhancement methods. Extended inserts, surface, coiled or twisted tubes and additives could be utilized in passive methods to enhance the efficiency while electrostatic fields, surface vibration, injection and suction could be considered in active method to increase the performance of heat exchanger. Shah and Sekulic (2007) studied temperature profiles in two different heat exchangers, and highlighted two major disadvantages in the parallel flow design. The large temperature difference at the ends caused large thermal stresses. The contraction of the construction materials and opposing expansion due to different fluid temperatures could eventually result in material failure. Besides, the temperature of the cold fluid exiting the heat exchanger remains lower than the lowest temperature of the hot fluid, and this can be improved by optimizing the operating parameters.

A counter flow heat exchanger has the most effective flow patterns (Bergman and Incropera, 2011). Moreover, it has the lowest heat exchanger surface area because of the highest log mean temperature differences (LMTD). Based on comparable conditions, more heat is transferred in a counter flow than in a parallel flow heat exchanger. Ghias et al. (2016) stated that LMTD method could be used to determine the overall heat transfer coefficient U from experimental values using inlet and outlet temperatures and the fluid flow rates. With the inlet temperatures and U known, this method is not useful in predicting the outlet temperatures. In this case, the effectiveness NTU method is convenient to predict the outlet temperature. Moran et al. (2003) discussed that the convective heat transfer coefficient relied on fluid properties, flow geometry, and the flow rate. It is appropriate to describe this dependence using several dimensionless numbers called the Reynold number, which determines the flow regime. Naphon (2007) stated that heat transfer occured because of conduction between the layers of the fluid under laminar regime. In the transition and turbulent regimes, heat transfer happens mainly by forced convection. Higher turbulence could result in enhanced heat transfer.

There are various types of shell and tube heat exchangers (Kakac et al., 2012). Baffles in the shell are used to enhance the rate of heat transfer. The quantity of shell side and the tube side flow arrangement would depend on the pressure drop, heat duty, fouling factor, manufacturing techniques, cost, corrosion control and the cleaning purpose. Lebele-Alawa and Egwanwo (2012) performed extensive numerical analysis on heat transfer in the shell and tube heat exchanger using basic governing equation and based on three parameters: outlet temperature, the heat exchanger effectiveness and heat transfers coefficient. Helical baffle was used to enhance the heat transfer rate as it forced the shell side to access the plug flow condition attaining higher efficiency. The types of baffles also affect the efficiency of the heat exchangers. Kirubadurai et al. (2014) found that modified baffle had a higher enhancement performance than segmental baffle.

Kondhalkar and Kapatkat (2012) carried out experimental study on the performance analysis of a spiral tube heat exchanger. They found that spiral tube heat exchanger had 15-20% lower cost than shell and tube type heat exchanger. The low velocity with more turbulence could reduce fouling and increase the heat transfer rate in the spiral tube heat exchanger, making it preferable than other types of heat exchanger. Hossain et al. (2012) analyzed the heat transfer coefficient and effectiveness of the spiral coil heat exchanger operating with water. It was found that heat transfer rate increased nearly with straight line with increased Reynolds number, and this was considered acceptable for the spiral coil heat exchanger. In addition, Nusselt number also increased with increasing Reynolds number for all three cold water flow rates tested. The work of Guha and Unde (2014) also showed the attainment of these requirements in the most optimal way along with achieving safety, operability, maintainability, sustainability and profitability. Bhavsar et al. (2013) focused on performance analysis of spiral tube heat exchanger and stated that heat transfer with a spiral tube was higher than that of a straight tube.

In counter flow arrangement, centrifugal force created by the spiral shape enhances the heat transfer of both tube and shell sides. Shirgire and Kumar (2013) also found that the heat transfer in helical coil heat exchanger was higher than that of straight tube heat exchanger due to compactness. Reddy et al. (2016) discussed that the contact time had influence on the heat transfer rate of heat exchangers. Shell and tube heat exchanger have a low contact time compared to other types of heat exchangers. As a result, increasing contact time would play a role in enhancing heat transfer rate. The sudden pressure drop is one of the major problems encountered in heat exchanger, affecting the efficiency and the overall heat transfer coefficient of the heat exchangers significantly. The objective of this study was, therefore, to investigate and compare the overall heat transfer coefficient between a single-pass shell and tube and spiral coil heat exchangers under co-current and counter current flows. The effects of process variables, such as temperature distribution, overall heat transfer coefficient, effectiveness, flow rate and pressure drop through hot testing under balanced flow condition are also investigated, which would benefit in designing and optimizing the performance of the two heat exchangers.

EXPERIMENTAL SET UP AND TECHNIQUES

Figure 1 shows the experimental setup utilized for the current study. It consists of storage tank, centrifugal pump, digital hydraulic pressure gauge, and infrared temperature sensor. Heater and pump were mounted so that hot water can be pumped and passes through the Shell. The electric heater of capacity 2 kW mounted in the hot reservoir was used to raise the temperature of water to 60-70^{o}C prior to pumping to the heat exchanger. The valves mounted at different locaton in the setup were used to have cocurrent and counter current flows in the two heat exchangers. Dimensions of the shell and tube and spiral coil are depicted in Table 1 and 2, respectively. The heat exchanger shell is made of IS2606 steel and has an internal diameter of 100 mm and thickness of 10 mm whereas the spiral coils has internal and outside diameters of 34 mm and 44 mm, respectively.

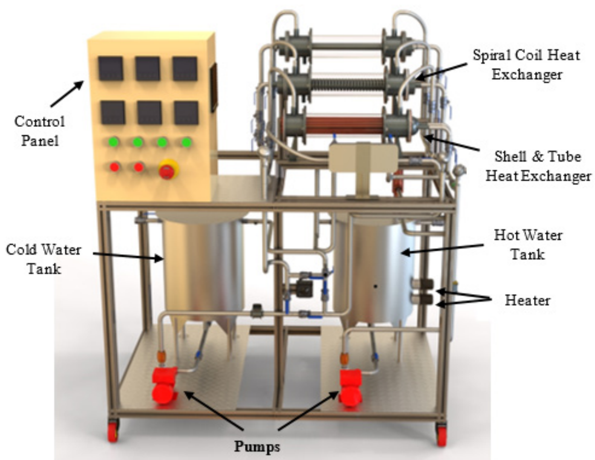


Figure 1. Experimental set-up

Table 1
Dimensions of shell and tube

Item	Dimension
Tube O.D. (do)	6 mm
Tube I.D. (di)	4 mm
Tube Length (L)	Approx. 440 mm at one pass
Tube Count (Nt)	16 (Two passes)
Tube Pitch (pt)	18 mm
Tube arrangement	Triangle
Shell O.D.	110 mm
Shell I.D. (Ds)	100 mm
Baffle Count	5
Baffle Distance (lB)	50 mm
Material of Construction	Stainless Steel/ Copper/ Borosilicate

Table 2
Dimensions of spiral coil

Item	Dimensions
Coil Tubing O.D.	9.53 mm
Coil Tubing I.D.	7.05 mm
Coil Length (L)	5.00 m
Coil I.D.	34 mm
Coil O.D.	44 mm
Shell O.D.	110 mm
Shell I.D. (Ds)	100 mm
Material of Construction	Stainless Steel/ Borosilicate

The mathematical equations were used to analyze the raw experimental data to evaluate the heat transfer rate, heat lost, logarithmic mean temperature difference and overall heat transfer coefficient. Heat absorbed or released was calculated by using the following expression:

$$Q = m\ Cp\ \Delta\ T \tag{1}$$

Where,
Q is heat transfer rate for hot or cold water (w/m^{2}K)
m is mass flow rate (kg/s)
Cp is specific heat of water (J/kg.K)
ΔT is temperature difference (^{o}C)

Heat lost was calculated as follows:

$$\text{Heat lost} = \text{Qhot} - \text{Qcold} \tag{2}$$

Efficiency was then calculated as:

$$Efficiency = \left(\frac{Qcold}{Qhot}\right) X\ 100\% \tag{3}$$

Overall heat transfer coefficient was calculated as follows:

$$U = \frac{Q}{A\ \Delta T_{lm}} \tag{4}$$

Where,
U = overall heat transfer coefficient
Q = heat transfer rate for hot or cold water
A = total contact area
ΔT_{lm}= logarithmic mean temperature difference (LMTD)

LMTD was calculated as follows:

$$\Delta T_{lm} = \frac{\Delta T1 - \Delta T2}{\ln(\Delta T1/\Delta T2)} \tag{5}$$

The temperature gradient for co-current flow was calculated as in Equations (6) and (7):

$$T_{hot\ in} - T_{cold\ in} = \Delta T_1 \tag{6}$$
$$T_{hot\ out} - T_{cold\ out} = \Delta T_2 \tag{7}$$

The temperature gradient for counter-current flow was given as:

$$T_{hot\ in} - T_{cold\ out} = \Delta T_1 \tag{8}$$
$$T_{hot\ out} - T_{cold\ in} = \Delta T_2 \tag{9}$$

The heat transfer effectiveness, ε, was calculated as follows:

$$\varepsilon = \frac{Q}{Q_{max}} \tag{10}$$

Where,

Q is the actual heat transfer rate

Q_{max} is the maximum possible heat transfer rate

The maximum possible heat transfer rate was calculated as follows:

$$Q_{max} = C_{min}\,(T_{hot\ in} - T_{cold\ in}) \tag{11}$$

C_{min} is the smaller heat capacity rate of hot water (C_{hot}) and cold water (C_{cold}).

RESULTS AND DISCUSSION

Comparative analysis between Shell tube heat exchanger and Spiral coil heat exchangers Figure 2 and 3 show the overall heat transfer coefficients of the spiral coil and shell and tube heat exchangers with co-current and counter current flow patterns, respectively. As the mass flow rate through the spiral coil and shell tube heat exchangers increased the overall heat transfer coefficient also increased. The overall heat transfer coefficient was much higher for the spiral coil heat exchanger than that of shell tube heat exchangers. The effectiveness of heat exchanger was greatly affected by hot water mass flow rate and cold water flow rates. When hot water mass flow rate increased, the effectiveness decreased. Spiral coil counter current flow was most effective in all these conditions and shell tube parallel flow heat exchanger was least effective.

Pressure drops in the spiral coil and shell and tube heat exchangers for the co-current and counter current flows are depicted in Figure 4 and 5, respectively. Pressure drop in the

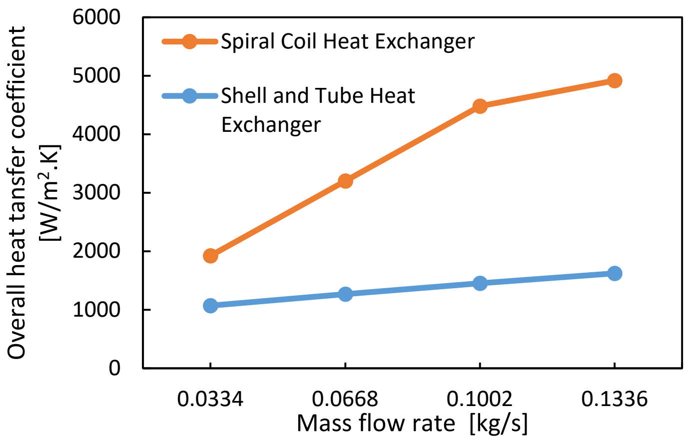


Figure 2. Comparison of overall heat transfer coefficient between the Spiral coil and Shell tube heat exchangers (co-current)

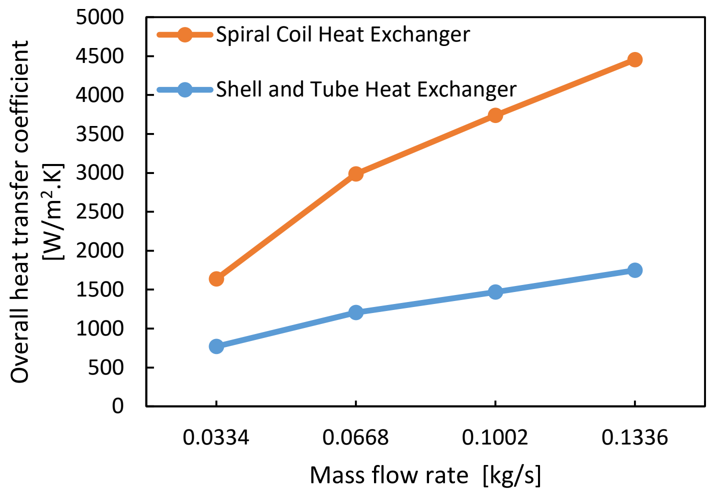


Figure 3. Overall heat transfer coefficient in the Spiral coil and Shell tube heat exchanger (counter current)

spiral coil was higher than that of the shell tube heat exchanger. Moreover, the spiral coil heat exchanger occupies less space (Andrzejczyk & Muszynski, 2018).

Comparative Analysis between Co-current Flow and Counter Current Flow

Figure 6 shows the difference in overall heat transfer coefficient between co current and counter current flows for the shell and tube heat exchanger. The overall heat transfer coefficient increased with an increase in the mass flow rate. However, the heat transferred for counter current flow was higher compared to the co-current flow with maximum overall heat transfer coefficient of 1750.34 W/m^2.k calculated in a laminar regime.

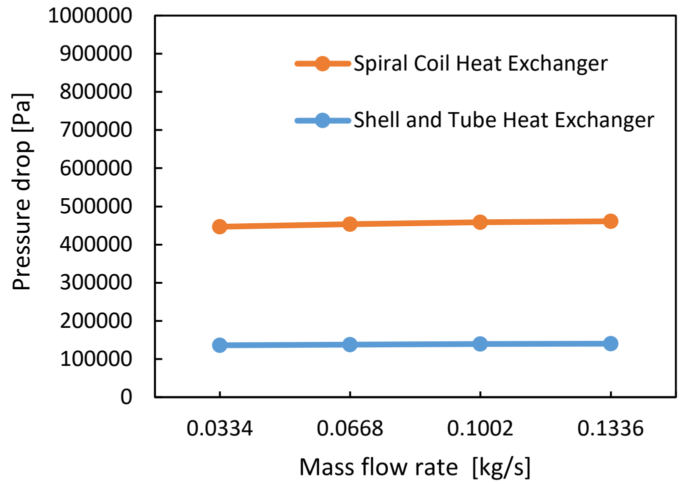


Figure 4. Comparison of pressure drop between the Spiral coil and Shell tube heat exchangers (co-current)

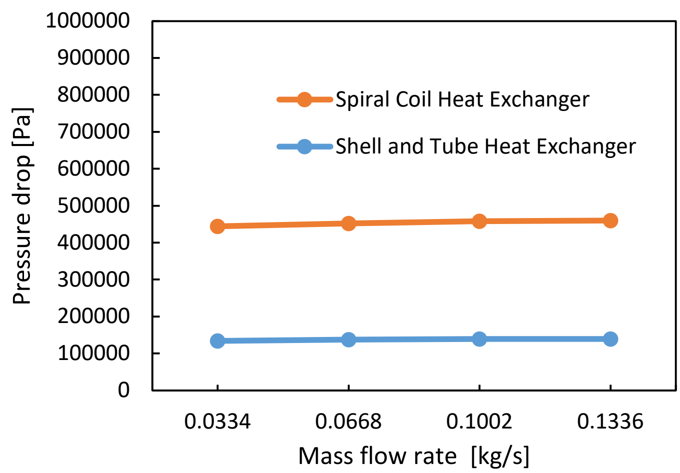


Figure 5. Pressure drop in the Spiral coil and Shell and tube heat exchangers (counter current)

Figure 7 shows the overall heat transfer coefficient between co current and counter current flow for the spiral coil heat exchanger. The overall heat transfer coefficient increased with an increase in the mass flow rate (Tapre & Kaware, 2015). However, the amount of the heat transfer for counter current flow was greater than that of the co-current flow with the maximum coefficient of 3295.67 W/m^2.k at a mass flow rate of 0.1336 kg/s.

Figure 8 shows pressure drop in the co-current and counter current flows of the shell and tube heat exchanger. It can be shown that the pressure drop in the shell and tube heat exchager for counter current flow increased with the mass flow rate. The amount of pressure

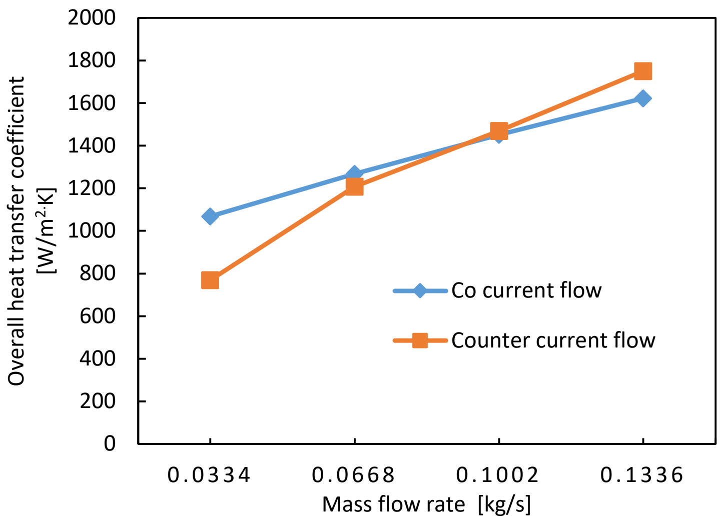


Figure 6. Comparison of overall heat transfer coefficient between co-current and counter current flows in the shell and tube heat exchanger

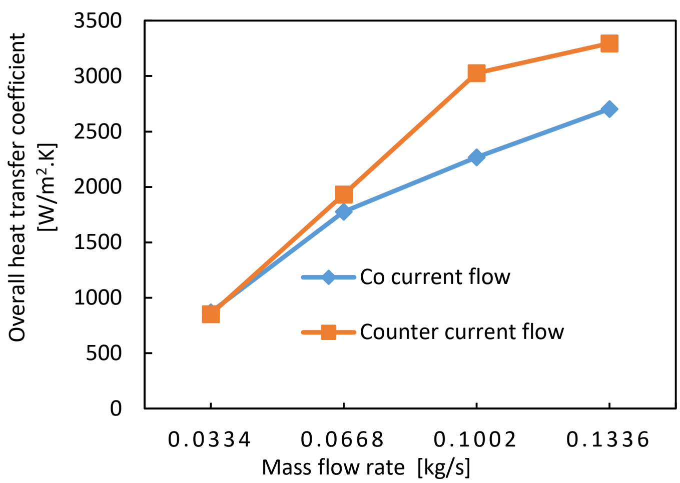


Figure 7. Overall heat transfer coefficient between co-current and counter current flows in the spiral coil heat exchanger

drop for the shell and tube in co-current flow also increased with the increase in mass flow rate and this was smaller than that of counter current flow.

Figure 9 shows pressure drop in cocurrent and counter current flows in the spiral coil heat exchanger. It can be seen that pressure drop in counter current flow of the shell and tube heat exchanger increased linearly with the mass flow rate. The pressure drop for the spiral coil in cocurrent flow also increased with an increase in mass flow rate.

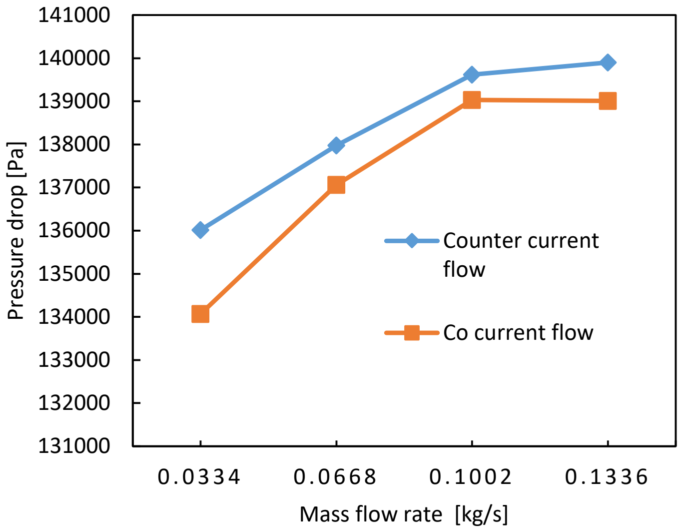


Figure 8. Pressure drop in the co-current and counter current flows in shell and tube heat exchanger

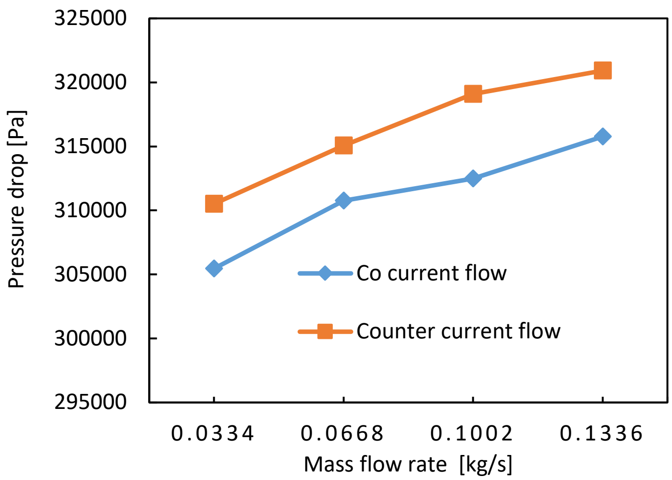


Figure 9. Pressure drop in the co-current and counter current flows in the spiral coil heat exchanger

Figure 10 shows the relationship between overall heat transfer coefficient against Reynold number for counter current and co-current shell and tube heat exchanger. It can be clearly seen that the overall heat transfer coefficient increased with an increase in the cold flow rates. Both flows were found laminar based on the Reynold number. Due to laminar flow, the velocity of the fluid flow through the shell is small and therefore the interaction among fluid particles is also low. When the velocity is low, the heat transfer among fluid particles occurs very slowly. This is why the overall heat transfer coefficient becomes quite similar at different Reynolds number.

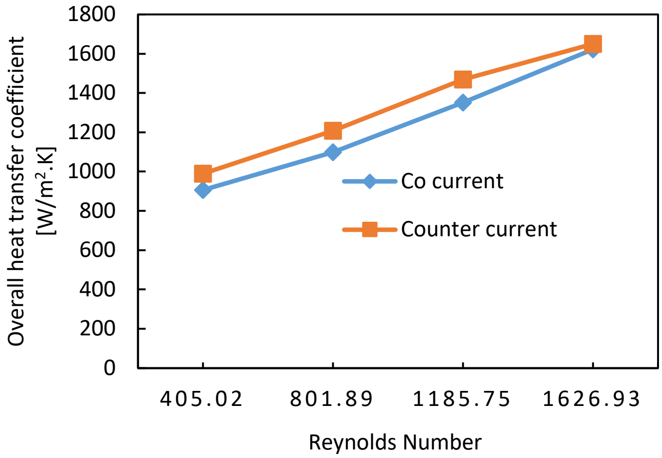


Figure 10. Overall heat transfer coefficient against Reynolds number for the shell and tube heat exchanger

Figure 11 shows the relationship between overall heat transfer coefficient, U and Reynold number for counter current and co-current spiral coil heat exchanger. It can be clearly seen that overall heat transfer coefficient increased with Reynolds number for both flow patterns. Besides, the overall heat transfer coefficients for counter current and co current flow were also quite similar since it works with the same variation of cold water flow rate. Overall heat transfer coefficient is also dependent on convective heat transfer coefficient, so an increase in Reynolds number resulted in higher heat transfer rates. It also shows that overall heat transfer coefficient for counter current is slightly higher than that of co-current flow patterns.

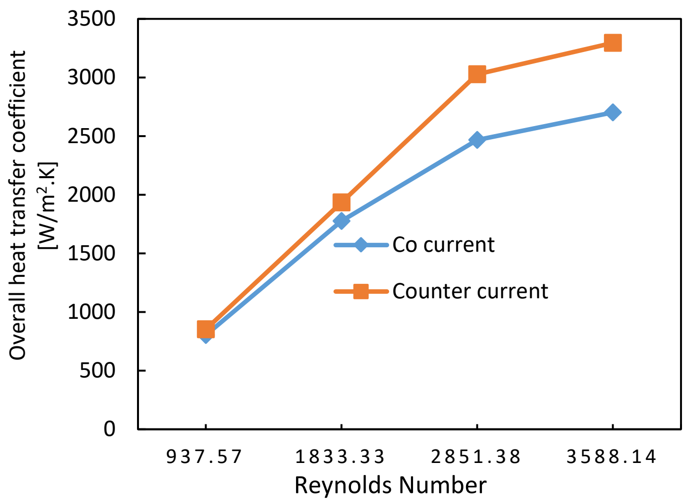


Figure 11. Overall heat transfer coefficient against Reynold number for the spiral coil heat exchanger

CONCLUSION

This paper presents a study on overall heat transfer coefficient of a single-pass shell and tube and spiral coil heat exchangers for better understanding of heat transfer coefficient and effectiveness of the two heat exchangers. The experiments were conducted on the shell and tube and spiral coil heat exchangers with different mass flow rates in cocurrent and counter current flow patterns. A comparison was made between cocurrent and counter current flows in the shell and tube and spiral coil heat exchangers with respect to the performance characteristics, such as overall heat transfer coefficient, heat transfer rate, temperature distributions, temperature difference and pressure drops. The overall heat transfer coefficient increased with an increase in mass flow rate. An increased mass flow rate induced an increase in pressure drop as expected. The counter current flow was more efficient compared to the co-current flow with enhanced overall heat transfer coefficient. Further research needs to be performed that include the experiment at low temperatures in order to check the performance of the present heat exchanger for various applications. Hence, the findings of this study would benefit the operator to design and optimize the best performance of heat exchangers.

ACKNOWLEDGEMENTS

The authors would like to thank Universiti Kuala Lumpur (UniKL) for the experimental facilities provided.

REFERENCES

Andrzejczyk, R., & Muszynski, T. (2018). An experimental investigation on the effect of new continuous core-baffle geometry on the mixed convection heat transfer in shell and coil heat exchanger. *Applied Thermal Engineering*, *136*, 237-251.

Bergman, T. L., & Incropera, F. P. (2011). *Fundamentals of heat and mass transfer*. Hoboken, NJ: John Wiley & Sons.

Bhavsar, J. J., Matawala, V. K., & Dixit, S. B. (2013). Performance analysis of spiral tube heat exchanger. *International Journal of Engineering and Advanced Technology, 2*(4), 684-687.

Faramarz S. R. (2013). Comparative study between helical coil and straight tube heat exchanger. *IOSR Journal of Mechanical and Civil Engineering (IOSR-JMCE), 8*(2), 55-59.

Ghias, A. S. A., Ananth, S. V., Anand, M. D., & Devadhas, G. G. (2016). Experimental study of thermal performance of coil in shell heat exchanger. *Indian Journal of Science and Technology, 9*(13), 1-17.

Guha, P., & Unde, V. (2014). Mathematical modeling of spiral heat exchanger. *International Journal of Engineering Research, 3*(4), 226-229.

Hossain, M. A., Onaka, Y., & Miyara, A. (2012). Experimental study on condensation heat transfer and pressure drop in horizontal smooth tube for R1234ze (E), R32 and R410A. *International Journal of Refrigeration, 35*(4), 927-938.

Kakac, S., Liu, H., & Pramuanjaroenkij, A. (2012). *Heat exchangers: selection, rating, and thermal design*. Boca Raton: CRC Press.

Kirubadurai, B., Rajasekaran, R., Kanagaraj, K., & Selvan, P. (2014). Heat transfer enhancement of shell and tube heat exchanger. *International Journal for Research in Applied Science and Engineering Technology, 2*(8), 309-318.

Kondhalkar, G. E., & Kapatkat, V. N. (2012). Performance analysis of spiral tube heat exchanger used in oil extraction system. *International Journal of Modern Engineering Research, 2*(3), 930-936.

Lebele-Alawa, B. T., & Egwanwo, V. (2012). Numerical analysis of the heat transfer in heat exchangers. *International Journal of Applied Science and Technology, 2*(4), 60-64.

Ma'arof, M. I. N., Chala, G. T., Husain, H., & Mohamed, M. S. (2019). Influence of fins designs, geometries and conditions on the performance of a plate-fin heat exchanger-experimental perspective. *Journal of Mechanical Engineering and Sciences, 13*(1), 4368-4379.

Mohanraj, M., Jayaraj, S., & Muraleedharan, C. (2015). Applications of artificial neural networks for thermal analysis of heat exchangers-a review. *International Journal of Thermal Sciences, 90*, 150-172.

Moran, M. J., Shapiro, H. N., Munson, B. R., & DeWitt, D. P. (2003). *Introduction to thermal systems engineering: Thermodynamics, fluid mechanics, and heat transfer*. New York: John Wiley & Sons.

Naphon, P. (2007). Thermal performance and pressure drop of the helical coil heat exchangers with and without helically crimped fins. *International Communications in Heat and Mass Transfer, 34*, 321-330.

Reddy, K., Rajasekaran, R., & Kanagaraj A. (2016). A review of heat transfer studies for shell and tube heat exchangers. *International Journal of Scientific Development and Research (IJSDR), 1*(5), 68-73.

Seyed, F. R. (2013). Comparative study between helical coil and straight tube heat exchanger. *IOSR Journal of Mechanical and Civil Engineering, 8*(2), 55-59.

Shah, R. K., & Sekulić, D. P. (2007). *Fundamentals of heat exchanger design*. Hoboken, NJ: John Wiley & Sons.

Shirgire, N., & Kumar, P. V. (2013). Review on comparative study between helical coil and straight tube heat exchanger. *IOSR Journal of Mechanical and Civil Engineering, 8*(2), 55-59.

Tapre, R. W., & Kaware, D. J. P. (2015). Review on heat transfer in spiral heat exchanger. *International Journal of Scientific and Research Publications, 5*(6), 2278-0181.

Zohuri, B. (2017). Heat exchangers. In B. Zohuri (Ed.), *Physics of cryogenics: An ultralow temperature phenomenon* (pp. 300-330). Amsterdam, Elsevier.

Pertanika J. Sci. & Technol. 27 (4): 2079 - 2107 (2019)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Optimization of Glass-beads Water Slurry flow Characteristics using Taguchi's Decision-making Technique coupled with Genetic Algorithm

Shofique Uddin Ahmed[1]*, Rajesh Arora[1] and Ranjana Arora[2]

[1]*Department of Mechanical Engineering, Amity University, Haryana 122413, India*
[2]*Department of Renewable Energy, Amity University, Haryana 122413, India*

ABSTRACT

This current study represents a numerical computational fluid dynamics (CFD) study of Glass beads-water slurry flow and aims for the optimization of various flow parameters adopting a new approach combining Taguchi's method and genetic algorithm (GA) for a smooth and better flow of the slurry. Three flow parameters viz. inlet mean velocity (V_m), volume fraction (C_{vf}) and particle size (D_p) with three levels were considered for the simulation and two responses vs. particle flow velocity (V_f) and pressure drop ($\frac{\Delta P}{L}$) were analysed to achieve a better understanding of the complex multiphase Glass-beads Water Slurry flow. Taguchi decision making technique and Genetic algorithm were exploited in order to optimize the particle flow velocity and pressure drop. A mathematical model was developed in order to forecast the responses analytically. Analysis of variance (ANOVA) was selected for finding the contribution of the selected flow parameters on particle flow velocity and pressure drop. The designed experiments were simulated numerically using analysis systems (ANSYS) software package by adopting Eulerian two-phase approach along with RNG K-ε model. Finally the optimized responses along with percentage reduction in pressure drop were reported and validated with the experimental data. This study optimized the pressure drop and particle flow velocity for three different input factors which would be helpful for the slurry-transporting industries, Oil and Gas industries and thermal power plants for optimal pipe design as well as to initiate a smooth slurry flow through horizontal pipeline.

Keywords: ANOVA, concentration distribution, genetic algorithm, pressure drop, regression analysis, slurry flow, Taguchi's method, velocity distribution

ARTICLE INFO

Article history:
Received: 20 September 2018
Accepted: 23 April 2019
Published: 21 October 2019

E-mail addresses:
uddinshofique1991@gmail.com (Shofique Uddin Ahmed)
rajesharora1219@gmail.com (Rajesh Arora)
ranjana1219@rediffmail.com (Ranjana Arora)
* Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

INTRODUCTION

Carrying solid materials through pipeline over a long distance has been imminent practice over the years in wide range of industries including oil and gas, manufacturing and production industries, food processing industries, city municipality, and pharmaceutical industries. Slurry flow is a process where solid particles are blended properly with fluid and is allowed to flow through pipelines over a long distance. The durability of this mode of solid material transport is that it is cost efficient as well as reduces air pollution, environmental hazards, and road traffic. Slurry flow is a complex multiphase flow problem. Flow parameters and flow patterns of slurry flow exhibit a complex behaviour and hence it is necessary to have detailed information about the slurry flow parameters and slurry flow patterns for a better understanding of the slurry flow process. Many classifications of slurry flow is available in the literature among which the most common classification includes homogenous flow, heterogeneous flow, stationary and moving bed flow and saltation flow regime. Among all these flow regime heterogeneous flow regime is the most occurring flow regime which shows a heterogeneous distribution of the solid particles within the fluid.

Computational fluid dynamics (CFD) is a sophisticated platform that allows the researchers to solve a wide range of complex flow problems at low cost and with greater ease. Many multiphase flow problems including slurry flow problems can be numerically analysed using this CFD technique which would have been very difficult to analyse experimentally. Hence the demand of CFD analysis is increasing day by day and researchers are adopting this cost efficient technique for their study. CFD provides exclusive idea about the variation of different flow variables inside the domain of the flow which plays a vital role in determining the behaviour of the flow process.

Taguchi's method is a sophisticated design technique for designing the experiments with least variance. It uses the concept of orthogonal array for arranging the experiments in proper order as well as allowing an optimal arrangement of the control parameters. Orthogonal Array produces a number of unbiased experimental settings with minimum number of experiments. Some logarithmic functions of the desired output or response are used in Taguchi's approach known as signal to noise ratio (S/N ratio) which acts as a key function in optimization process. Furthermore, Taguchi's technique aids to determine the desirable results and analysis of data. Generally, the larger the better, nominal the best and lower the better S/N ratios are used. S/N ratio considers the mean as well as the variability. S/N ratio is the ratio of mean aka signal to the standard deviation aka noise. Taguchi's technique allows an improvement of the productivity by reducing the influence of deviation without affecting the causes.

Analysis of Variance (ANOVA) is a statistical approach to developing impetus comparisons between two or more means; the yield results using this method can be tested to predict whenever powerful symbolic correlation exists between the variables. ANOVA is an assemblage of statistical models and their combined measure where the inspected variance of a special variable is divided into components imputable to various provenances of variation. ANOVA is efficacious in confrontation of two, three or more means.

Researches in this area began at 3rd decades of 20th century. O'Brien (1933) and Rouse (1937) performed an experimental investigation on open channel slurry flow containing very low volume fraction and they studied the concentration distribution of fully suspended solid particles. Toda et al. (1973) conducted an experimental comprehensive study on two phase solid liquid slurry flow where they investigated the mean particle flow velocity for pipe bends, vertical and horizontal pipelines. They concluded that the particle velocities remained distributed in a wide range as well as the flow pattern of each particle exhibiting a different path. Turian and Yuan (1977) developed a pressure drop correlation for a multiphase slurry flow considering homogenous, heterogeneous, saltation flow and stationary bed flow. Their developed correlation showed an improvement in predicting pressure drop of the slurry flow. Oroskar and Turian (1980) developed a correlation implementing the energy used to suspend the particles to predict the critical velocity of slurry flow. Their proposed correlation was improved version of the previously available correlation for predicting the critical velocity. Roco and Shook (1983) developed a sequence of numerical equations for the prediction of velocity and concentration distribution of a quasi-uniform slurry flow considering different particle size. In their study explicit algorithm was used for determining numerical solutions of different flow patterns and the effects of solid particle size were determined. Doron et al. (1987) conducted an investigation of coarse particles slurry flow through horizontal tubes, they proposed a physical model to predict the pressure drop and flow patterns of the slurry flow and their proposed model was compared with their experimental results which showed a good consistency with the experimental results. Gillies et al. (1991) studied the slurry flow of gravel, sand and coarse coal particles and they proposed a mathematical model for predicting the head losses for these slurry flow. The contact load offering sliding friction at the wall of the pipe was found to be the objective function of their proposed model while this model was limited to the slurry flows containing less than 35% volume fraction of the particles. Gillies et al. (2004) conducted an experimental investigation of sand water slurry flow containing different particle size through horizontal pipeline having 103mm pipe diameter. In their study various flow variables such as *in situ* velocity and solid concentration and their effects on pressure drop were obtained and it was found that the friction loss showed a reduced effect at higher velocities. They also proposed a mathematical relation for predicting the friction loss at higher velocities and they used their proposed relation to rectify the model to determine slurry flow friction. Kaushal and Tomita (2007) studied the slurry flow of glass beads and water considering two particle sizes viz. 125μm and 440μm through a horizontal pipeline. They concentrated their study on determining the near wall lift of the solid particles using γ-ray densitometer. From their study it was concluded that for the finer particles the highest concentration zone was near the pipe bottom while the highest concentration zone of coarser particles were established away from the base of the pipe. Lahiri and Ghanta (2010) developed a CFD generalized model for predicting the concentration distribution and pressure drop of a water- glass beads slurry flow adopting Eulerian two phase approach and RNG K-ε approach. The predicted pressure drop and concentration distributions were validated with the experimental data and were in an excellent consistency. Their proposed model showed an improvement for the prediction of flow parameters of slurry flow than the previously available models. Chandel et al. (2010) studied the rheological characteristics and pressure drop of a slurry flow containing fly ash and bottom ash above 50% concentration by weight using an experimental setup. Finally they determined the pressure gradient of slurry flow in straight pipes at high concentration using their rheological data and the pressure gradient was compared with the experimental data for the validation of their study. There are very few researches on optimization of flow problem using Taguchi's technique. Taguchi's technique helps us optimizing flow conditions selecting a set of optimal flow parameters that fits right for the optimum flow of the fluid through pipeline. Some of the researches in this field include the work of Wu et al. (2008); who reduced the micro bubble drag in a turbulent channel flow using Taguchi's robust design approach. In their study they looked into the impact of several controllable flow variables such as microbubble size, area of air injection and amount of air injected on microbubble drag reduction. From their study they found out that for optimum parametric conditions the effect of microbubble drag can be reduced upto 21.6%. Kotcioglu et al. (2013) conducted a comprehensive study for the selection of optimum design parameters in a rectangular duct heat exchanger using Taguchi's technique. They investigated the effect of various design variables such as pressure drop, flow velocity, Reynolds number, inclination angle of winglets, ratio of the winglets length to the duct channel length and ratio of the duct channel width to height for the performance parameter response of friction factor and Nusselt number.

This present study aims for the optimization of flow conditions for the smooth flow of water- glass beads slurry using a new approach combining Taguchi's method with Genetic Algorithm. Eulerian two phase approach along with RNG K-ε model was adopted for modelling the multiphase flow and turbulence phase of the flow respectively.

This present study represents CFD numerical analysis of glass beads-water slurry flow through a 54.9mm diameter and 4m length. In depth analysis of velocity profiles and concentration profile were carried out in order to achieve a better understanding of the complex multiphase slurry flow process. Three different flow parameters viz. inlet mean velocity, solid volume fraction and particle size were analyzed and their effects on pressure drop were determined. Experiments were designed in an unbiased fashion using Taguchi's L_9 Orthogonal Array. A logarithmic function S/N ratio was introduced in order to optimize individually the particle flow velocity and pressure drop for three different flow conditions viz. inlet mean flow velocity, solid volumetric concentration and particle size. In addition a numerical model was developed using regression analysis to envisage the responses analytically. A new approach for multi-objective optimization was implemented exploiting the advantages of Genetic Algorithm and a specific combination of flow conditions was obtained which is capable of optimizing both the responses simultaneously. The contributions of these flow conditions on pressure drop and particle flow velocity were determined using Analysis of Variance statistical analysis tool. A range of particle sizes viz. 90µm, 125 µm and 150µm for inlet mean velocity of 3m/s, 4m/s and 5m/s with solid volume fraction of 30%, 40% and 50% were considered as the simulation parameters.(Flow Parameters). Finally the optimized responses obtained from GA analysis were compared with the simulated results as well as the simulated results of pressure drop were validated with the experimental results of pressure drop available in the literature of Kaushal and Tomita (2007). In addition the percentage reduction of pressure drop using this new strategy was reported in this present study.

METHODS

Model Description

Slurry flow problem is a multiphase flow problem. There are many mathematical models available foe modelling the multiphase flow in the commercial CFD software package ANSYS. Slurry flow consists of continuous liquid phase and discrete solid phase which cannot be considered as diffused dilute system. Here in this study Eulerian two-phase flow model had been selected for modelling the multiphase flow and RNG K-epsilon approach was selected for modelling the turbulence phase of the flow. The solid particles were assumed to be mono-dispersed. In other words the solid phase was assumed to behave like liquid because of its dispersed nature. One of the advantages of selecting Eulerian two-phase model is that the Eulerian model satisfies both the momentum and mass equation for the solid and fluid phase individually while at the same time the mixture model fails to predict the pressure drop correctly. Moreover, granular version of Eulerian two-phase model is preferred for modelling the slurry flow because unlike the non-granular version it takes into account of collision and friction among the solid particles which is an important factor that has to be considered while modelling the slurry flow.

Eulerian Two-Phase Model

This model comprises most complex equations of multiphase modelling in FLUENT software. In Eulerian model, the slurry is supposed to comprise solid and liquid phases. The concentrations of the two phases are assumed to be α_s and α_f for $\alpha_s + \alpha_f = 1$. In the present study, granular flows had been undertaken comprising fluid/solid intermixing whose modelling could be achieved by appropriate constraints of interphase exchange and pressure coefficients. Moreover, the characteristics of the flow had been derived with the applications of kinetic theory. The various forces that act on a single solid particle are given as follows (Kaushal et al., 2012).

1. Static pressure gradient ΔP_s.
2. Inertial force caused by particles of solid pressure gradient ΔP_s.
3. Forces due to the difference in velocities of two phases, $K_{sf}\left(\vec{v}_s - \vec{v}_f\right)$
4. Viscous and body forces, $\nabla.\overline{\overline{\tau_f}}$ and $\rho\vec{g}$, where $\overline{\overline{\tau_f}}$ represents the stress tensor of fluid, ρ denotes the mass density and g is gravitational acceleration.
5. Lift/virtual mass forces. The coefficient of virtual mass/ lift forces, C_L/C_{vm} are assumed to be 0.5.
6. The particles in the analysis are assumed to be fluid in nature.

In this paper, an efficient Eulerian-Eulerian multi-phase model is adopted for numerical simulation. The following governing equations are used for the turbulent flow of fly ash particles in the Newtonian fluid.

Governing Equations

Continuity Equation

$$\nabla.(\alpha_t \rho_t \vec{v}_t) = 0, \; t \text{ being either } \textit{solid} \text{ or } \textit{fluid.} \quad (1)$$

Momentum equation for fluid phase

$$\nabla.\alpha_f \rho_f \vec{v}_f \vec{v}_f = -\alpha_f \nabla P + \nabla.\bar{\bar{\tau}}_f + \alpha_f \rho_f \vec{g} + K_{sf}\left(\vec{v}_s - \vec{v}_f\right) + C_{vm}\alpha_f \rho_f\left(\vec{v}_s \nabla \vec{v}_s - \vec{v}_f \nabla \vec{v}_f\right)$$
$$+ C_L \alpha_s \rho_f (\vec{v}_f - \vec{v}_s) \times \nabla \times \vec{v}_f \quad (2)$$

Momentum equation for solid phase

$$\nabla.\alpha_s \rho_s \vec{v}_s \vec{v}_s = -\alpha_s \nabla P - \nabla P_s + \nabla.\bar{\bar{\tau}}_s + \alpha_s \rho_f \vec{g} + K_{sf}\left(\vec{v}_f - \vec{v}_s\right) + C_{vm}\alpha_s \rho_f\left(\vec{v}_f \nabla \vec{v}_f - \vec{v}_s \nabla \vec{v}_s\right)$$
$$+ C_L \alpha_s \rho_f (\vec{v}_s - \vec{v}_f) \times \nabla \times \vec{v}_f \quad (3)$$

$\bar{\bar{\tau}}_f$ and $\bar{\bar{\tau}}_s$ being the stress tensor for the fluid and solid phase respectively.

$$\bar{\bar{\tau}}_f = \alpha_f \mu_f\left(\nabla \vec{v}_f + \nabla \vec{v}_f^{tr}\right) \quad (4)$$

$$\bar{\bar{\tau}}_s = \alpha_s \mu_s (\nabla \vec{v}_s + \nabla \vec{v}_s^{tr}) + \alpha_s (\lambda_3 - \tfrac{2}{3}\mu_s) \nabla \vec{v}_s \bar{\bar{I}} \quad (5)$$

The bulk viscosity of solid can be donated by λ_s and is represented as

$$\lambda_s = \frac{4}{3}\alpha_s \rho_s d_s g_{o,ss}(1 + e_{ss})\left(\frac{\Theta_s}{\pi}\right)^{\frac{1}{2}} \tag{6}$$

The turbulence model is solved by using RNG K-eplison approach with other additional conditions causing interfacial turbulent momentum transfer. The Reynolds stress tensor for the fluid phase is given by

$$\bar{\bar{\tau}}_{t,f} = -\frac{2}{3}\left(\rho_f k_f + \mu_{t,f}\nabla\vec{v}_f\right)\bar{\bar{I}} + \mu_{t,f}\left(\nabla\vec{v}_f + \nabla\vec{v}_f^{tr}\right) \tag{7}$$

Here $\mu_{t,f}$ is the fluid viscosity corresponding to the turbulence. An analytical differential interrelationship for turbulent viscosity is provided with RNG K-epsilon approach for modelling the flow at lower Reynolds number. However, for the sake of the present study which is based on high Reynolds number the analytical correlation converts to

$$\mu_{t,f} = \rho_f C_\mu \frac{k_f^2}{\varepsilon_f} \text{ with } C_\mu = 0.09 \tag{8}$$

The prediction of turbulent kinetic energy k_f and turbulent energy dissipation rate ε_f is by Standard K-epsilon method and RNG K-epsilon method is almost alike. Standard and RNG K-epsilon approach differ in such a fashion that RNG K-epsilon approach contains a supplementary expression in the epsilon equation

$$R_\varepsilon = \frac{C_\mu \rho \eta^3 (1-\eta/\eta_0)\varepsilon^2}{(1+\beta\eta^3)k} \tag{9}$$

Where $\eta = Sk/\varepsilon$, $\eta_0 = 4.38$, $\beta = 0.012$, the constant parameters are taken as $C_{\mu m} = 0.0845$, $C_{1\varepsilon} = 1.42$, $C_{2\varepsilon} = 1.68$, $C_{3\varepsilon=1.3}$ $\sigma_k = 0.75$, $\sigma_\varepsilon = 1.2$

Solution Strategy

In this current study the Navier-Stokes governing equations along with their closure parameters are solved using ANSYS FLUENT 14.0. Momentum and mass equations are solved by control volume finite difference approach. Turbulent kinetic energy, momentum equation and turbulent dissipation rate are discretized using second order upwind scheme while for discretizing the volume fraction first order upwind scheme is used in this study. For pressure and velocity coupling phase coupled SIMPLE algorithm is selected. Convergence of problem depends on the scaled residual. For this flow problem the residuals contains continuity, X-velocity, Y- velocity, Z-velocity for both the phases; k and ε for phase 1 and volume fraction for phase 2 which need to be converged at some specific region. Application of these schemes confirmed better stability, accuracy and convergence of the flow problem. Moreover decreasing the value of under relaxation factors ensures better convergence of the problem. URF (under relaxation factor) for volume fraction has been reduced to 0.3 from 0.5 and for momentum it has been reduced from 0.7 to 0.5.

Geometry and Mesh

A circular pipe of 54.9mm diameter and 4m length had been modelled in ANSYS WORKBENCH design modular. A non-uniform structured mesh with hexa core elements was selected for meshing the whole computational domain as shown in Figure 1. Seven layers of inflations were selected with 0.005mm size and growth rate 1.2 for refining the mesh near the wall of the pipe. This is important for achieving more precise results near the wall boundary.

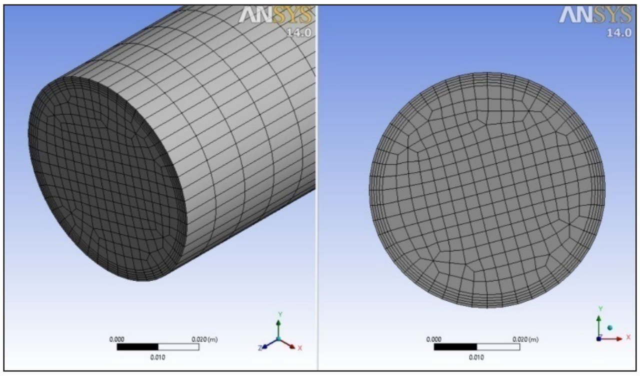


Figure 1. 3D and 2D view of the mesh selected for this study.

Grid Independency Test

Selection of proper grid size is highly recommended in most of the CFD numerical analysis problems. It provides better and enhanced simulation results. In this study, five mesh with different numbers of elements viz 0.95 Lakhs, 1.50 Lakhs, 2 Lakhs, 3.1 Lakhs and 3.89 Lakhs were introduced and simulations were conducted imposing same boundary conditions(V = 5m/s, C_{vf} = 0.3) for each of the mesh as shown in Figure 2. It had been found from the velocity plots for each of the mesh, that the velocity distribution of mesh having 2 Lakhs elements and 3.10 Lakhs elements were super imposing, hence the mesh with 2 Lakhs elements had been considered as the optimal mesh in this present study.

Boundary Conditions

Three boundary conditions had been imposed to the computational flow domain namely inlet boundary, wall boundary and outlet boundary. At inlet boundary, inlet mean flow velocity and solid volume fraction was introduced, while at wall boundary No-slip boundary conditions had been imposed. The outlet boundary had been treated as pressure outlet for solid as well as fluid phase.

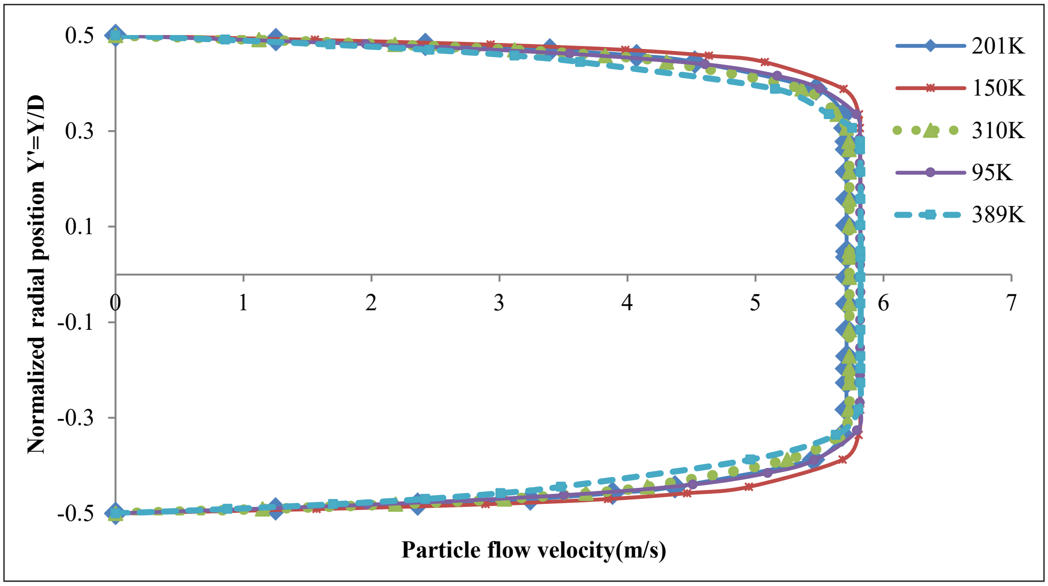


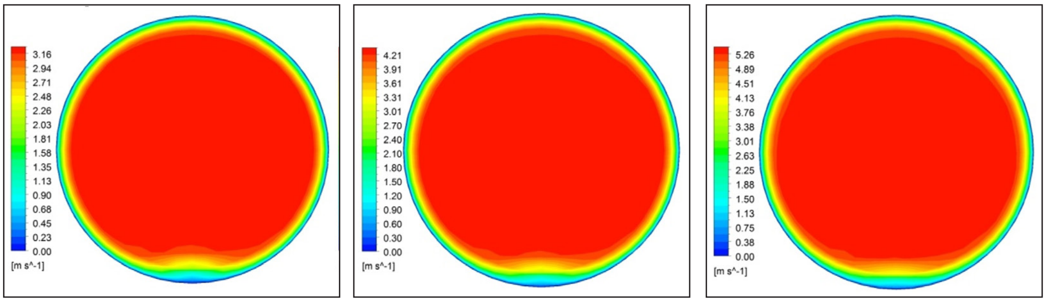
Figure 2. Plot of Grid Independency test

RESULTS AND DISCUSSION

Velocity Distribution

Figures 3 to 6 show the distribution of particle flow velocity at various volume fraction and particle size along the cross section of the pipe. The velocity contours are obtained at 3.9m length of the pipe from the inlet.

Figures 3 and 4 illustrate the particle flow velocity distribution of 90µm particle size and 30% solid volume fraction for different inlet mean velocity. From these figures it can be observed that at low inlet mean velocity (Figure 3a) the particle flow velocity distribution was asymmetrical at the base of the pipe, but this asymmetric nature reduced as the inlet mean velocity increased (Figure 3c).This is because of the reason that as the



(a) (b) (c)

Figure 3. Particle flow velocity distribution (V_f) of 90µm particle size and 30% solid volume fraction (C_{vf}) for different inlet mean velocity (V_m): (a) V=3m/s; (b) V=4m/s; (c) V=5m/s

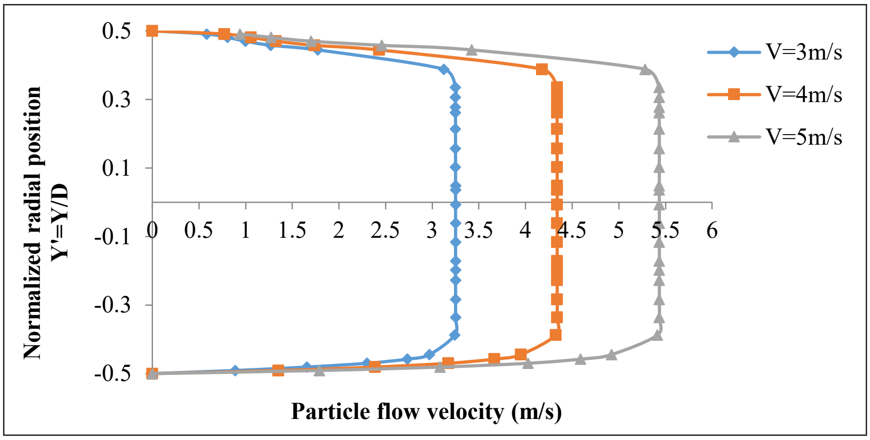


Figure 4. Plot of Particle flow velocity distribution (V_f) of 90µm particle size and 30% solid volume fraction (C_{vf}) for different inlet mean velocity (V_m)

inlet velocity becomes high, there is an increase in the flow turbulence, providing a perfect blending for the solid particles with the fluid. As a result, the solid particles became more distributed throughout the entire fluid region and hence, lesser number of solid particles could accumulate at the base of the pipe. Hence, the velocity distribution becomes more symmetrical at high inlet velocity (say V_m = 5m/s).

Figures 5 and 6 show the effect of solid volume fraction and particle size on particle flow velocity at a given inlet mean velocity. It can be noticed from Figure 6 that the velocity distribution was influenced by the solid particle size. As a result, the velocity profile became more distorted as the particles became larger in size from 90µm to 150µm. This may be the result of gathering of coarse particles at the base of the pipeline due to gravitational effect. Moreover, the degree of distortion of particle velocity is influenced by the solid volume fraction. For instance, at low volume fraction (C_{vf} = 0.3), the velocity profile is

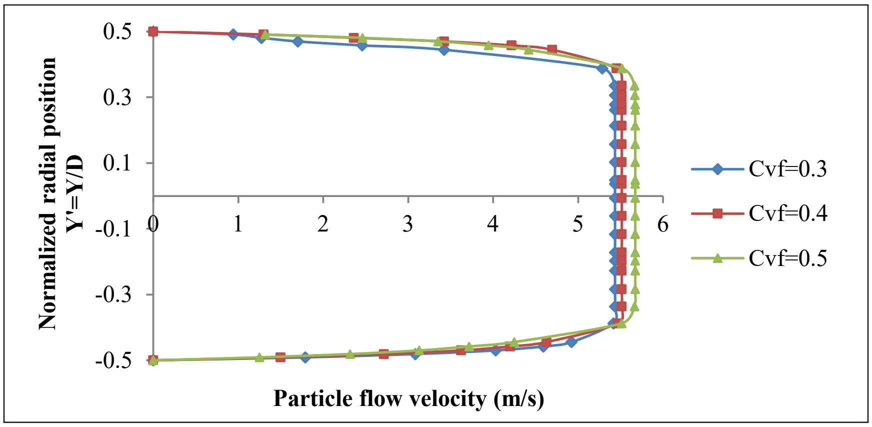


Figure 5. Plot of particle flow velocity distribution (V_f) at constant inlet mean velocity (V_m = 5 m/s) for different solid volume fraction (C_{vf}= 0.3, 0.4 and 0.5)

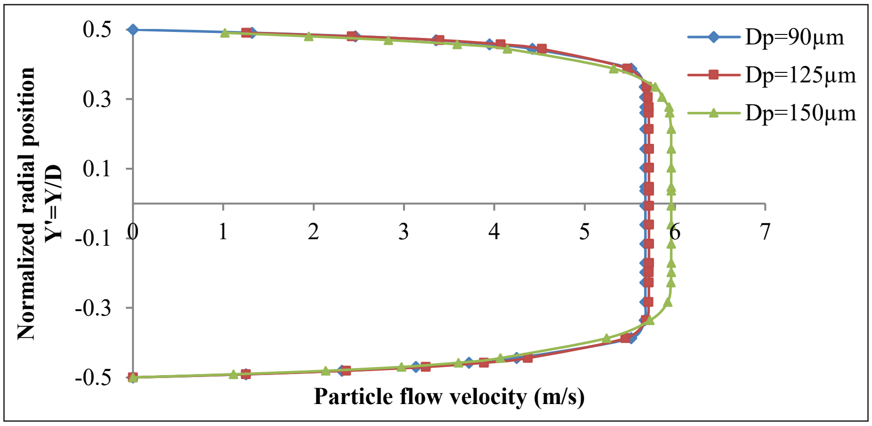


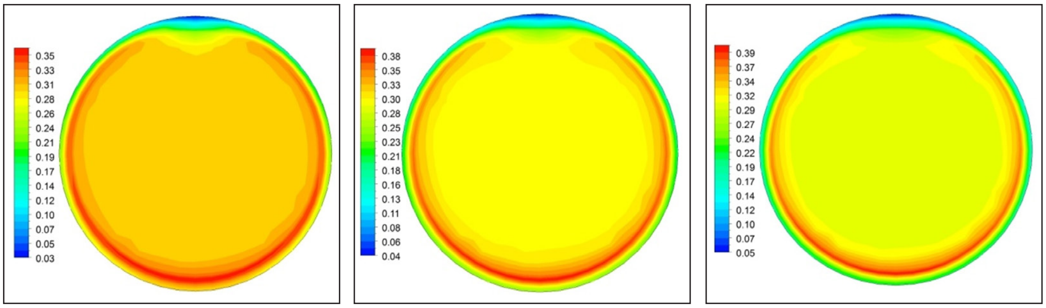
Figure 6. Plot of particle flow velocity distribution (V_s) at constant inlet mean velocity (V_m = 5m/s) at constant volume fraction (C_{vf}= 0.5) for different particle size D_p)

less distorted as well as the velocity is lower than that of high-volume fraction (C_{vf} = 0.5) (Figure 5). This might be the effect of increased turbulence energy at high volume fraction. In addition to this, it can be observed that the particle flow velocity decreases with the decrease in particle size for each solid volume fraction. As a result, smaller particles show more buoyancy effect which leads to symmetrical distribution and better blending of particles with the fluid that causes smoother flow of solid particles throughout the pipeline.

From this section it can be summarized that for obtaining maximum particle flow velocity a slurry flow with higher inlet mean velocity and high solid volume fraction as well as larger particle size is highly recommended.

Concentration Distribution

Figures 7 to 9 show the distribution of local concentration of solid particles along the vertical centreline of the pipe outlet, $Y' = Y / D$, Y is the vertical distance of the pipe from top to bottom and D is diameter of the pipe.



(a) (b) (c)

Figure 7. Concentration distribution of 90µm particle size and 30% solid volume fraction (C_{vf}) for different inlet mean velocity (V_m): (a) V=3m/s; (b) V=4m/s; (c) V=5m/s

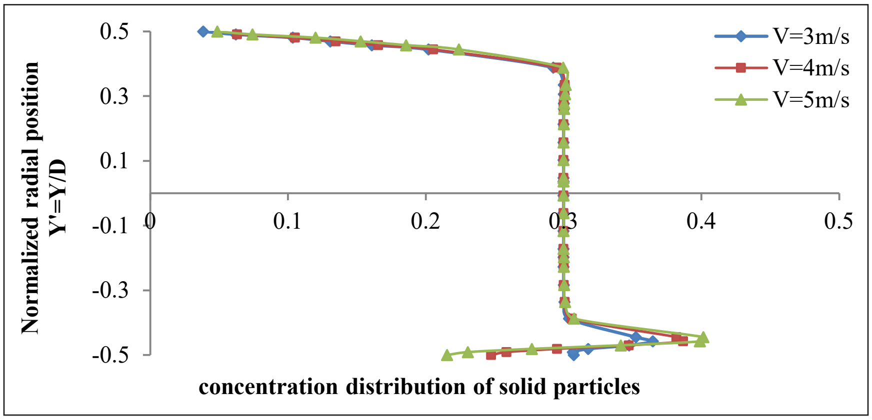


Figure 8. Plot of Concentration distribution of 90µm particle size and 30% solid volume fraction (C_{vf}) for different inlet mean velocity (V_m)

From Figures 7 to 8, it is noteworthy that the distribution of solid particles was asymmetrical across the vertical centreline of the pipe outlet, the high distribution zone being at the lower half of the pipe, where most of the solid particles seemed to accumulate at the base of the pipe. For low inlet mean velocity, the particles were accumulated at the base of the pipe (Figure 7a), but as the inlet mean velocity increased the solid particles became more buoyant and suspended in the fluid (Figure 7c) thereby providing their proper mixing with the fluid and ensuring a smooth slurry flow. This is because of the fact that when the inlet velocity increases, there is an increase in turbulence in the flow which is accountable for the mixing of solid particles with the fluid. Hence, at high inlet velocity (say 5m/s) the solid particles remains suspended in the fluid rather than settling down at the base of the pipe and hence, the concentration distribution becomes more distributed and symmetrical.

From Figure 9 it can be noticed that the distribution of solid particles was more homogenous when the particle size was small. It is also seen that the smallest particles (D_p= 90µm) being most buoyant as well as the distribution of particles in the fluid phase was homogenous. The particle concentration zone was distributed symmetrically over the vertical cross section at pipe outlet. However, as soon as the particle size became larger (from 90µm to 150µm) the concentration zone was shifted towards the lower portion of the pipe i.e. the base of the pipe cross section because of the gravitational effect, as a result the concentration zone became asymmetrical. As a result, the larger particles concentrate at the base of the pipe forming a stationary bed and refuse to travel further along with the fluid to the exit of the pipe. Further, as the particles become larger the thickness of the bed tends to increase and thereby reducing the effective flow area, which leads to poor productivity. The simulated results showed that the maximum concentration at the pipe bottom for 90µm is 0.35 whereas for 150µm particles it goes up to 0.41.

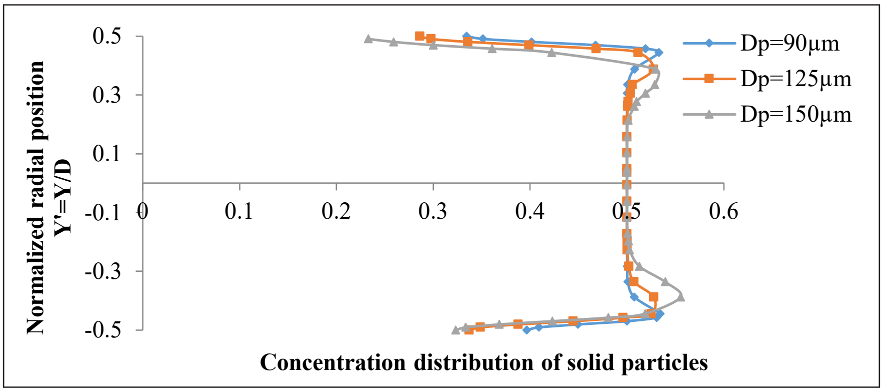


Figure 9. Plot of concentration distribution of solid particles at constant inlet mean velocity (V_m) and constant solid volume fraction (C_{vf}) for different particle size

Influence of Inlet Mean Velocity on Pressure Drop

Figure 10 shows the plot of effect of inlet mean velocity on pressure drop at different solid volume fraction for 125µm particle size.

From Figure 10 it is noteworthy that the pressure drop rose with the increase in inlet flow velocity. This is the result of increase in turbulence in the flow at high velocities which led to an increase in pressure drop. Moreover, as the volume fraction became high, the rate of rise in pressure drop also increased.

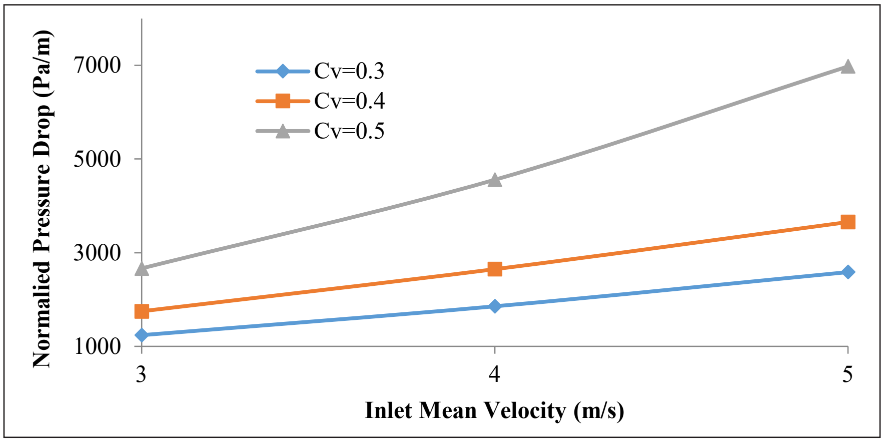


Figure 10. Plot of the effect of inlet mean velocity on pressure drop

Effect of Particle Size on Pressure Drop

Figure 11 shows the effect of particle size on pressure drop at different inlet mean velocity. It is noteworthy that there is a surge in pressure drop as the particles become larger. This happens because the larger particles tended to concentrate at the pipe bottom due to gravitational effect and thereby reducing the effective flow area which caused an increase in turbulence of the flow. The increase in rate of pressure drop at higher inlet mean velocities can also be observed from the figure. Moreover, it can be noticed that the solid volume fraction plays an important role during the change in pressure drop. The degree of rise in pressure drop is low for lower volume fraction because at low volume fraction the turbulence is lower but when the solid volume fraction increases there is an increase in turbulence and hence, the pressure drop increases abruptly.

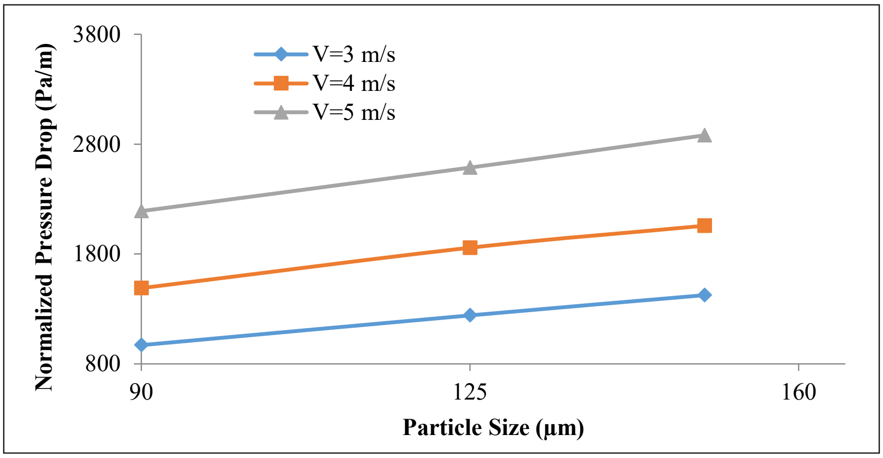


Figure 11. Plots of effect of particle size on pressure drop at constant volume fraction (C_{vf}= 0.3) for different inlet mean velocity

Effect of Volume Fraction on Pressure Drop

From Figure 12 it can be noticed that the pressure drop arose at higher volume fraction. At lower volume fraction (0.3) the pressure drop was lower and it rose as the volume fraction increased to 0.5. This is because of the fact that the turbulence of the flow becomes higher

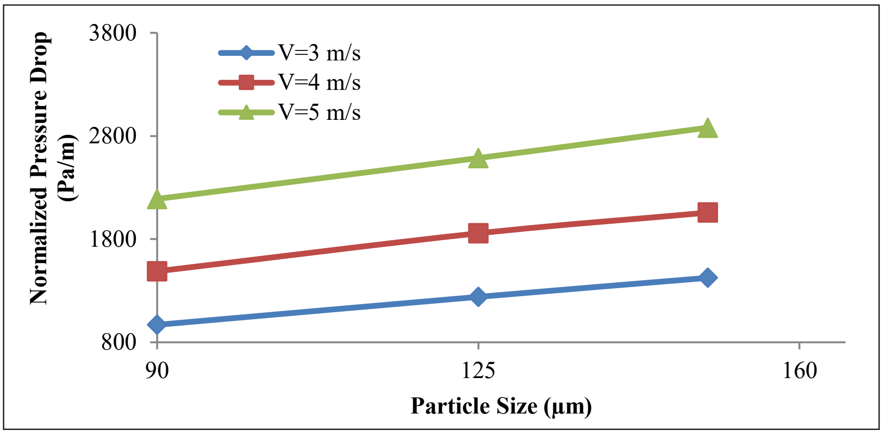


Figure 12. Effect of volume fraction (C_{vf}) on pressure drop at different particle size

with the increase in volume fraction thereby increasing the pressure drop. Furthermore, it was observed that with larger particle size the pressure drop also increased.

Prediction of Optimum Slurry Flow Conditions using Taguchi's Decision Making Technique

The aim of this section is to determine optimum flow condition for the maximum particle flow velocity and nominal pressure drop. When the particle flow velocity is high, the solid particles flow in a faster rate as well as the suspension of the solid particles also increases providing a complete mixing of solid particles with the fluid and ensuring a smooth slurry flow with minimal bed formation at the base of the pipe (Chandel et al., 2010). This will eventually increase the productivity. Furthermore, when the pressure drop is higher, there is an increase in frictional loss since pressure drop is directly proportional to frictional loss. Moreover, when the pressure drop is less, due to low turbulence, the solid particles tend to accumulate at the pipe bottom, hence creating a bed which reduces the effective flow area and hence productivity. Therefore, a nominal pressure drop is required for the smooth flow of slurry through the pipeline.

To determine the optimal flow condition Taguchi's L_9 orthogonal array was selected and inlet mean velocity, solid volume fraction and particle size were considered as input parameters (factors) of L_9 orthogonal array considering three levels for each of the factor. Since the desired output was maximum particle flow velocity, larger the better S/N ratio had been selected and can be represented by the following mathematical equation

$$S/N = -10 \times \log_{10} [(1/n) \times \mathrm{Sum}\ (1/y_i^2)] \qquad (10)$$

Here, n=Number of response and Y_i= particle flow velocity

Since Nominal pressure drop was the desired output, nominal is the best S/N ratio had been selected and can be represented by the following mathematical equation

$$\mathrm{Eta} = 10 \times \log_{10} (\mathrm{Mean}^2/\mathrm{Variance}) \qquad (11)$$

Table 1 represents selected factor and their levels where as L_9 orthogonal array of Taguchi's robust design approach is shown in Table 2.

Table 1
Selected factor and their levels

Factors	Level 1	Level 2	Level 3
Inlet mean velocity (V_m) m/s	3	4	5
Solid volume fraction (C_{vf})	0.3	0.4	0.5
Particle size (D_p) µm	90	125	150

Table 2
L_9 orthogonal array of Taguchi's robust design approach

Run	Factors		
	A	B	C
1	1	1	1
2	1	2	2
3	1	3	3
4	2	1	2
5	2	2	3
6	2	3	1
7	3	1	3
8	3	2	1
9	3	3	2

Design of Experiments

Table 3 shows the assignment table of L_9 orthogonal array adopted in this study whereas Response (Particle flow velocity) obtained from CFD simulated results and their corresponding S/N ratio is shown in Table 4.

Table 3
Design matrix of the experiments using Taguchi's L_9 orthogonal array

Experiment no.	Column		
	V_m (m/s)	C_{vf}	D_p (μm)
1	3	0.3	90
2	3	0.4	125
3	3	0.5	150
4	4	0.3	125
5	4	0.4	150
6	4	0.5	90
7	5	0.3	150
8	5	0.4	90
9	5	0.5	125

Results of particle flow velocity (response) and Pressure drop obtained from CFD simulation of the above experiments are tabulated in Table 4.

The following Figures 13 and 14, obtained from Minitab software shows the plots of mean S/N ratio vs. the flow parameters i.e. inlet mixture velocity, solid volume fraction and particle size.

For the analysis of the results obtained from the experiments designed by Taguchi's approach, S/N ratio analysis is used. Since "Larger the better" S/N ratio was selected for the analysis of particle flow velocity hence the highest mean of S/N ratio in Figure 14

Table 4
Response (Particle flow velocity and Pressure drop) obtained from CFD simulated results and their corresponding S/N ratio

Experiment no.	Particle flow velocity (m/s)	Pressure drop (Pa/m)	S/N ratio for particle flow velocity	Mean for pressure drop
1	3.249	969.16	10.2350	969.16
2	3.418	1750.285	10.6754	1750.28
3	3.617	3092.183	11.1670	3092.18
4	4.403	1856.318	12.8750	1856.32
5	4.2134	2973.085	13.2438	2973.08
6	4.543	3573.058	13.1469	3573.06
7	5.511	2880.548	14.8246	2880.55
8	5.513	3046.658	14.8278	3046.66
9	5.712	6974.703	15.1358	6974.70

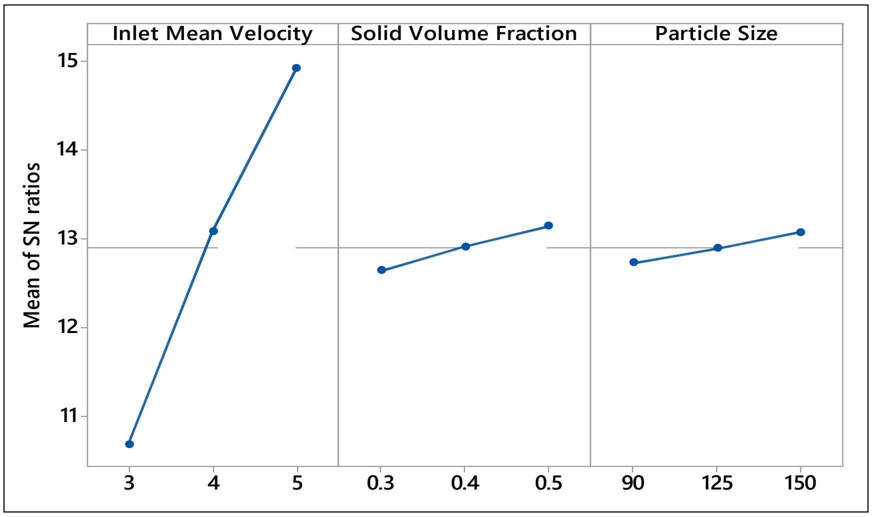


Figure 13. Influence of different flow parameters at different levels on S/N ratio of maximum particle flow velocity.

corresponds to the optimal level of input variables. In other words from the S/N analysis for maximum particle velocity the following conclusions can be drawn

- The inlet mean velocity should be 5m/s
- Solid volume fraction should be 0.5
- Particle size should be 150 µm

ANOVA Results

ANOVA was performed for determining the contributions of the flow parameters on particle flow velocity.

Table 5
Analysis of Variance results for particle flow velocity

Source	DOF	Seq SS	Contribution	Adj SS	Adj MS	F-Value	P-Value
Inlet mean velocity	2	6.93825	98.38%	6.93825	3.46913	5628.65	0.000
Solid volume fraction	2	0.08379	1.19%	0.08379	0.04190	67.98	0.014
Particle size	2	0.02907	0.41%	0.02907	0.01453	23.58	0.041
Error	2	0.00123	0.02%	0.00123	0.00062		
Total	8	7.05234	100.00%				

Table 5 shows that the highest contribution on particle flow velocity is inlet mean velocity (98.38%) followed by solid volume fraction (1.19%) and particle size (0.02%).

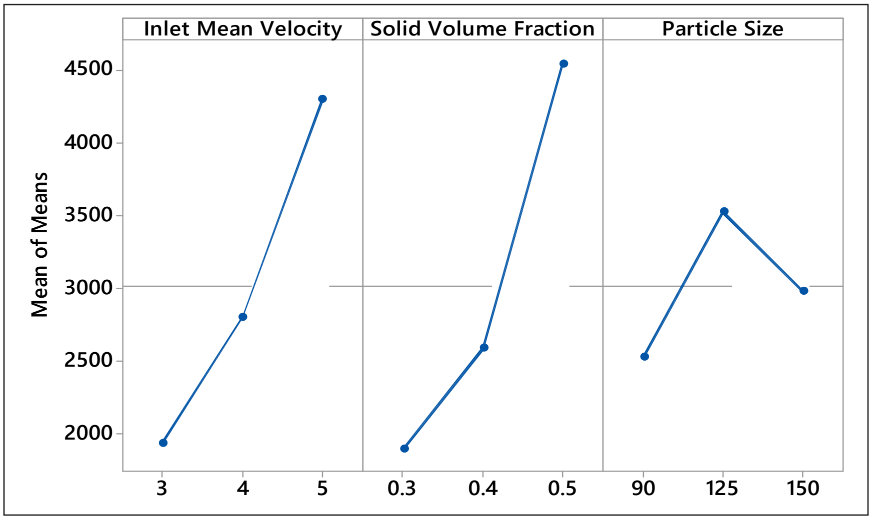


Figure 14. Influence of various flow parameters at different levels on Mean of nominal pressure drop.

Since "Nominal is best" S/N ratio has been adopted for the analysis of pressure drop, hence the nominal value (middle value) of the plot (Figure 15) represents the optimal level of the input parameters. In other words, the S/N analysis for nominal pressure drop shows the following results:

- Inlet mean velocity should be 4m/s
- Solid volume fraction should be 0.4
- Particle size should be 125 µm.

ANOVA was performed to determine the contributions of flow parameters on pressure drop.

From Table 6, it can be noticed the highest contribution on pressure drop was of solid volume fraction (48.41%) followed by inlet mean velocity (39.6%) and particle size (5.31%).

Table 6
Analysis of Variance results for pressure drop

Source	DF	Seq SS	Contribution	Adj SS	Adj MS	F-Value	P-Value
Inlet Mean Velocity	2	8581057	37.08%	8581057	4290528	4.84	0.171
Solid Volume Fraction	2	11295895	48.81%	11295895	5647948	6.38	0.136
Particle Size	2	1496750	6.47%	1496750	748375	0.85	0.542
Error	2	1771234	7.65%	1771234	885617		
Total	8	23144936	100.00%				

The optimal combination of input variables obtained from S/N ratio analysis is enlisted in Table 7.

Table 7
Optimal Parameter combinations obtained from S/N ratio analysis for both the responses.

Optimal Parameter combination				Responses	
Combinations	Inlet Mean Velocity (m/s)	Solid Volume Fraction	Particle size (μm)	Particle flow velocity	Pressure drop
I	**5**	**5**	**150**	**5.9581**	6832.87
II	**4**	**4**	**150**	4.21	**2973.085**

Note. Bold values represent the optimal parameters and optimal responses

From Table 7 it can be observed that both the responses cannot be optimized simultaneously adopting Taguchi's S/N ratio analysis. Hence to overcome this problem a more sophisticated analysis is required. In this study to optimize both the responses Genetic Algorithm has been adopted.

Development of Numerical Models

Regression Analysis. Linear Regression model has been adopted to develop a correlation between multiple variables when there is a correlation among the dependent variables and independent variables. In this study V_m and $\frac{\Delta P}{L}$ are considered as dependent variables while V_m, C_{vf} and D_p are considered as independent variables. The general model equation for multivariable linear regressions is:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots \dots \dots \dots \dots \dots + \beta_p x_p + \varepsilon \tag{11}$$

Where x is the independent variable and y is the dependent variable. β's are coefficient (regression parameters) and ε is the residue.

From this general model equation mathematical equations are developed for prediction of Particle flow velocity and pressure drop for various control parameters.

The Regression Equation for Particle flow velocity V_f and pressure drop $\frac{\Delta P}{L}$ can be given as follows

$$V_f = 1.07533V_m + 1.1817C_{vf} + 0.002307D_p - 0.5481 \quad (12)$$

With $R^2 = 99.98\%$

And

$$\frac{\Delta P}{L} = 1182V_m + 13223C_{vf} + 0.89D_p - 8084 \quad (13)$$

With $R^2 = 83.39\%$

The R^2 value signifies the goodness of the regression equation compared to the simulated/actual value. R^2 Value for pressure drop achieved from linear regression analysis is not good enough to be used as predictive model equation. Hence this regression equation has been modified adopting the 2nd order multi-regression analysis. The obtained equation is represented as follows

$$\frac{\Delta P}{L} = 582.473403V_m - 11293.3395C_{vf} - 30.3219765D_p - 2865.31176V_m \times C_{vf} + 337.913053C_v \times D_p + 36.2469524V_m \times D_p - 273.246176{V_m}^2 - 1.02945031{D_p}^2 - 195.066387 \quad (14)$$

With $R^2 = 99.9991\%$

Figures 15 and 16 shows the goodness of the fit for predictive model equation of particle flow velocity and pressure drop respectively.

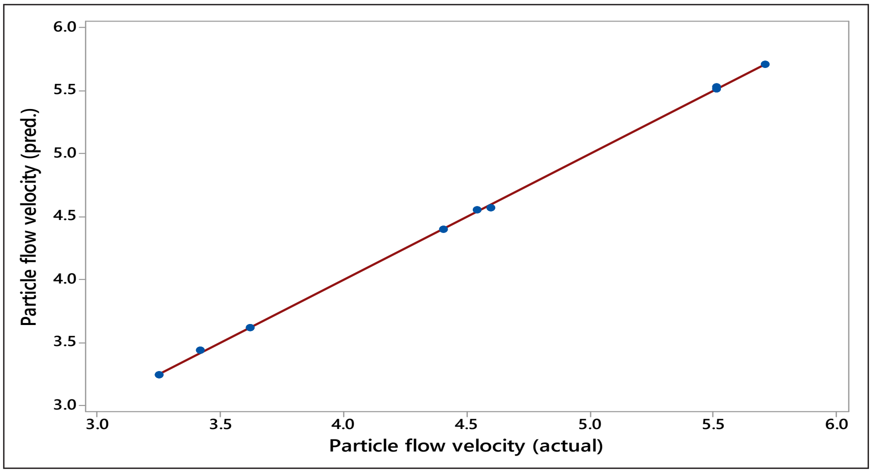


Figure 15. Regression plot of particle flow velocity (predicted) vs. Particle flow velocity (simulated)

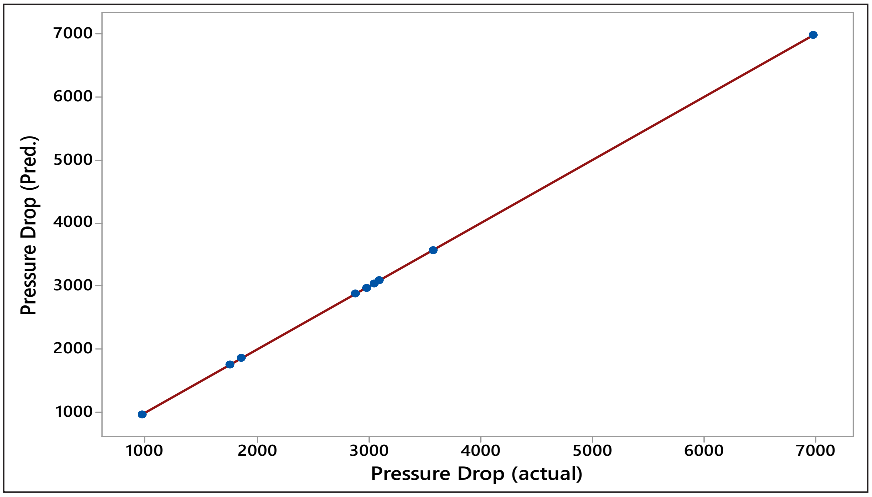


Figure 16. Regression plot of particle flow velocity (predicted) vs. Particle flow velocity (simulated)

From Figures 15 and 16 it can be noticed that the responses obtained from regression model equations for particle flow velocity and pressure drop shows a well aligned nature with the simulation data. Hence a predictive model consisting of these two equations can predict the particle flow velocity and pressure drop with minimal error (of the order of 0.001%).

Validation of the Developed Mathematical Model

The outcome obtained from the predictive mathematical model show a good reliability with the simulated results. In addition, to establish the robustness and reliability of this proposed model these outcomes have been compared with the experimental data taken from the work of Kaushal and Tomita (2007) and recorded in Table 8.

Table 8
Comparison of simulated and experimental responses with the proposed predictive mathematical model

	Flow Parameter combination			Responses	
	V_m (m/s)	C_{vf}	D_p (μm)	V_f (m/s)	$\frac{\Delta P}{L}$ (Pa/m)
Predicted by Proposed model	3	0.4	125	3.43902	1750.28
	4	0.3	125	4.39619	1856.32
	5	0.5	125	5.70786	6974.70
Experimental (Kaushal and Tomita ,2007)	3	0.4	125	-	1913.6
	4	0.3	125	-	2088.69
	5	0.5	125	-	6887.58
Simulated	3	0.4	125	3.418	1750.23
	4	0.3	125	4.403	1856.27
	5	0.5	125	5.712	6974.65

From Table 8 it can be observed that the predicted outcomes of the mathematical model are reliable and show excellent consistency with the experimental data. This model can predict the pressure drop data with an average error of 6.95% when compared with experimental data.

Multi-Objective Optimization of Particle Flow Velocity and Pressure Drop using Genetic Algorithm

Genetic algorithm (GA) is an artificially intelligent evolutionary algorithm which is capable of optimizing single objective as well as multiple objectives (responses) for various control variables within a specific range. GA is often preferred for optimizing multi-objective problem because it yields more accurate and effective outcome compared to other evolutionary algorithm (Oktem, 2009). GA is inspired by the biological evaluation process: Darwin's theory of survival for the fittest. The optimization process is initiated generating random potential solutions known as *'chromosomes'*. The whole set of chromosomes create a "*population"*. Chromosomes evolve in the process of several iterations known as '*generation.*' New offspring are evaluated by the process of '*crossover'* and '*mutation'*. Crossover divides each chromosome into two parts and combines one half with other half of a different chromosome. While mutation flips over a single chromosome. Chromosomes are then evolved with the usage of specific fitness criteria and best chromosomes remains in the solution while the rest are discarded. This process keeps repeating until at least one chromosome is evolved which possesses the quality of best fitness and it is considered as optimal solution.

Finding a specific combination of process variables which would maximize the particle flow velocity and reduce the pressure drop simultaneously is necessary in order to initiate a smooth slurry flow through the selected pipe section which would yield maximum transport efficiency. The optimal combination of process variables obtained from S/N ratio analysis can minimize only one response at a time (Table 6). Specific combination of control factors can be achieved within the range of optimal control factors enlisted in Table 6 with the help of Multi-objective GA optimization technique.

Chosen Objectives

The objective functions chosen are as under:

$$V_f = 1.07533V_m + 1.1817C_{vf} + 0.002307D_p - 0.5481 \qquad (15)$$

$$\frac{\Delta P}{L} = 582.473403V_m - 11293.3395C_{vf} - 30.3219765D_p - 2865.31176V_m \times C_{vf} + 337.913053C_{vf} \times D_p + 36.2469524V_m \times D_p - 273.246176{V_m}^2 - 1.02945031{D_p}^2 - 195.066387 \qquad (16)$$

Design/Decision Variables

The design parameters chosen in this present work are as under:

- Inlet mean velocity V_m in m/s
- Solid Volume Fraction C_{vf}
- Particle size D_p in μm

Constraints

The ranges for these variables are given as under:

$$4\text{m/s} \leq \mathrm{V_m} \leq 5\text{m/s}$$

$$0.4 \leq \mathrm{C_{vf}} \leq 0.5$$

$$125\ \mu\text{m} \leq \mathrm{D_p} \leq 150\ \mu\text{m}$$

Fitness function was created from equation 16 and 17 using MATLAB and codes were written accordingly. The following genetic algorithm critical factors are used in this proposed research work

Population type: Double vector
Population size: 50
Creation function: feasible population
Tournament size: 2
Mutation: Adaptive feasible
Crossover: Constraint dependent (ratio 1.0)
Maximum number of generations: 1000

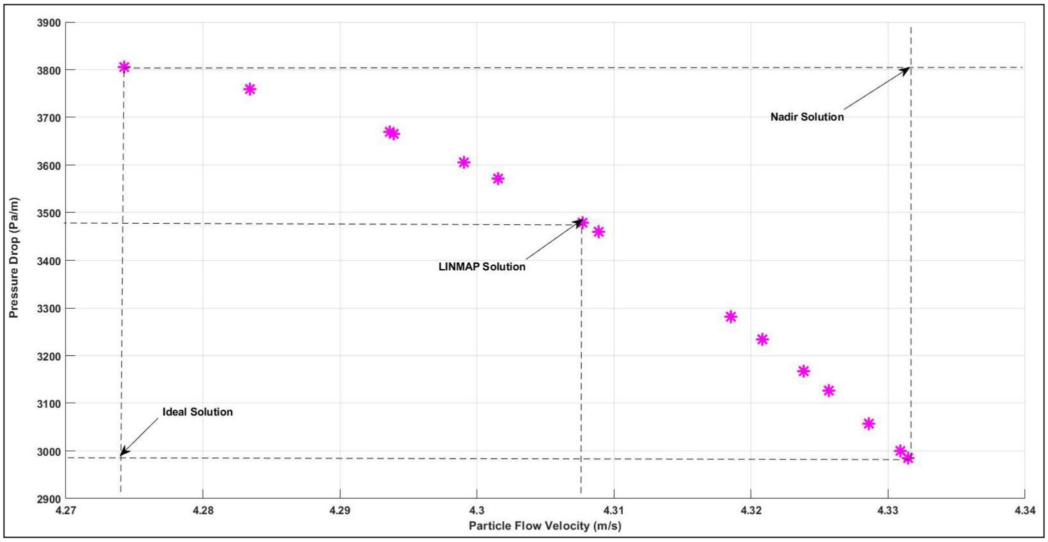


Figure 17. Plot of Pareto optimal solution for Particle flow velocity and Pressure drop obtained from GA optimization

In multi-objective optimization problem, a typical pareto frontier is attained and a final outcome from pareto set is obtained by suitable decision-making approach. It is presented in Figure 17. In the present study, the aim is to achieve minimum pressure drop for maximizing the particle flow velocity in order to encounter a smooth slurry flow process. All the values in the pareto front represent optimal values and hence, LINMAP method was employed for getting optimum solution of aforesaid objectives from pareto front. In case of LINMAP method the magnitude of distance between ideal/nadir points and the position of each individual was evaluated by distance matrices called Euclidian distance which is evaluated as below:

$$D_{i+} = \sqrt{\sum_{j=1}^{n} \left(F_{ij} - F_j^{ideal}\right)^2} \qquad (17)$$

LINMAP approach evaluates the minimum distance with respect to ideal value as follows:

$$i_f = i \in \min(D_{i+}) \qquad (18)$$

Employing this approach, the optimal value of Particle flow velocity (V_f) and pressure drop ($\frac{\Delta P}{L}$) are found to be 4.309 m/s and 3460.108 Pa/m respectively. A specific combination of process variables can be obtained at this optimal point as:

- V_m = 4 m/s
- C_{vf} = 0.4
- D_p = 140.002 µm

In other words, using a combination of inlet mean velocity of 4 m/s, solid volume fraction of 0.4 and particle size of 140.002 µm optimal particle flow velocity of 4.309 m/s and pressure drop of 3460.108 Pa/m can be acquired in a Glass-water slurry flow through 4 m long and 54.9 mm diameter horizontal pipe. This optimal condition can be used for optimal design of the pipeline and to encounter a smooth slurry flow through horizontal pipeline.

Validation of GA Optimized Results with Simulated Results

The simulated results show excellent consistency with the experimental data taken from the work of Kaushal and Tomita (2007). Hence the reliability of GA analysis has been carried out comparing the outcome obtained from GA analysis with the CFD simulated results. The specific combination of input parameters obtained from GA analysis (V_m = 4 m/s, C_{vf} = 0.4 and D_p = 140.002 µm) was used in CFD simulation and the outcomes are recorded in Table 8.

Table 9
Comparison of responses obtained from GA optimization strategy with CFD simulation results

	Optimal Input Parameters			Responses	
	Inlet Mean Velocity (m/s)	Solid Volume fraction	Particle Size (μm)	Particle Flow Velocity (m/s)	Pressure Drop (Pa/m)
GA Optimized	4	4	140.002	4.309	3460.108
Simulated	4	4	140.002	4.597	3392.5

The recorded data in Table 9 show that the GA optimized results are in excellent consistency with the CFD simulated results with mean percentage error of 4.13%. Hence this new strategy can be used to optimize multiple responses simultaneously in a complex multiphase slurry flow problem considering a range of flow parameters.

Reduction in Pressure Drop Implementing the GA Optimization Strategy

The optimal combination of flow parameters for individual optimization of particle flow velocity and pressure drop was attained from S/N ratio analysis while designing the experiments using Taguchi's method (Table 7). The combination is essential in order to have a primary idea about the ranges of flow parameters that would yield optimal responses in GA analysis. From Table 7 the mean optimal particle flow velocity and mean optimal pressure drop was calculated to be 5.084 m/s and 4902.9775 Pa/m. These responses had been compared with the outcome of GA analysis (Table 9) in order to find the % increase in particle flow velocity and % reduction in pressure drop as follows:

$$\% \text{ increase in Particle flow velocity} = \frac{\mathrm{V}_{f(\mathrm{Taguchi})} - \mathrm{V}_{\mathrm{f(GA)}}}{\mathrm{V}_{\mathrm{f(Taguchi)}}} \times 100\%$$

$$\frac{5.084 - 4.309}{5.084} \times 100\% = 15.24\%$$

$$\% \text{ Reduction in Pressure Drop} = \frac{\frac{\Delta P}{L}_{(Taguchi)} - \frac{\Delta P}{L}_{(GA)}}{\frac{\Delta P}{L}_{(Taguchi)}} \times 100\%$$

$$\frac{4902.9775 - 3460.108}{4902.9775} \times 100\% = 29.42\%$$

Hence after implementing the GA approach and using the specific flow conditions obtained from GA analysis 15.24% increase in particle flow velocity and 29.42% reduction in pressure drop can be observed over the responses obtained from Taguchi's method and S/N ratio analysis.

Validation of the Study

The aim of this section is to legitimize the simulated results of pressure drop obtained from CFD analysis with the experimental results of pressure drop available in the previous literature (Kaushal and Tomita, 2007). The Figures 18 to 20 show the plot of simulated pressure drop and the experimental pressure drop at 0.3, 0.4 and 0.5 volume fraction for a particle size of 125µm respectively.

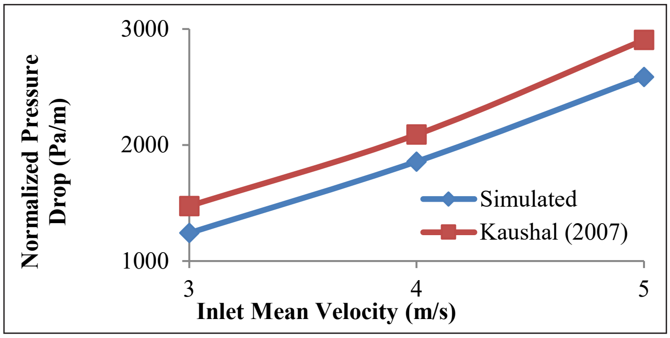


Figure 18. Comparison of simulated and experimental pressure drop at 0.3 solid volume fraction for 125µm particle size

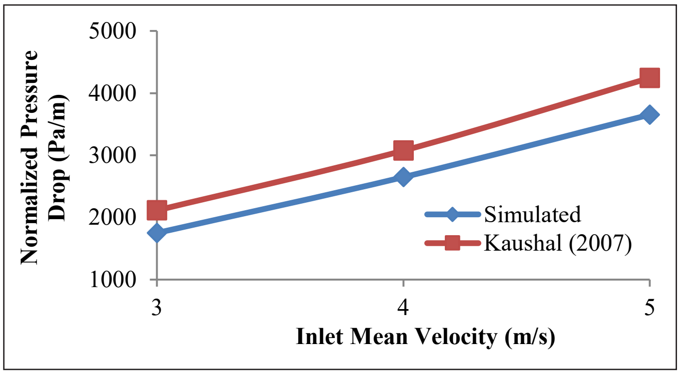


Figure 19. Comparison of simulated and experimental pressure drop at 0.4 solid volume fraction for 125µm particle size

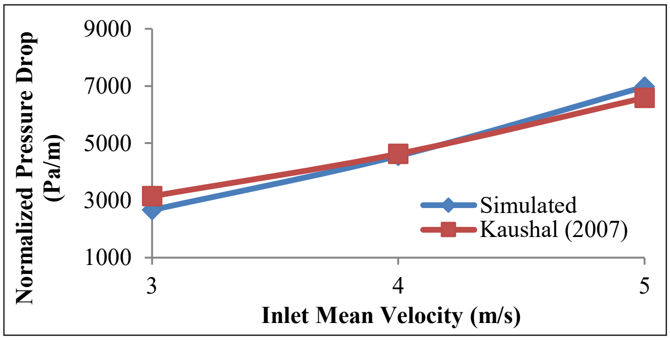


Figure 20. Comparison of simulated and experimental pressure drop at 0.5 solid volume fraction for 125µm particle size

CONCLUSION

This study represents an Eulerian two phase glass beads water slurry flow considering three different glass beads (specific gravity 2.47) particle size viz. 90µm, 125µm and 150µm through a horizontal pipe of 4m long and 54.9 mm diameter. The flow simulations are conducted at three different inlet mean velocities viz. 3m/s, 4m/s and 5m/s along with three different solid volume fraction viz. 0.3, 0.4 and 0.5 by volume. In this study the influence of various flow variables on pressure drop has been analysed as well as optimum flow conditions for smooth slurry flow has been analysed using Taguchi's decision-making technique and ANOVA. Multi-objective optimization of the responses have been undertaken in order to optimize both the responses simultaneously for a specific combination of flow parameters. Based on this study the following conclusions can be drawn:

- For all particle size the particle flow velocity distribution is asymmetrical at the lower part of the pipe at low inlet mean velocities. However, as the mean velocity increases the particle flow velocity distribution becomes less asymmetrical.
- At a constant inlet mean velocity the particle flow velocity increases as the solid volumetric concentration of the particles in the slurry becomes higher. Furthermore, the simulated results showed that the maximum concentration at the pipe bottom for 90µm is 0.35 whereas for 150µm particles it goes up to 0.41.
- The particle flow velocity is low for the smaller particle size and it increases as the particle size becomes larger.
- The solid particles are heterogeneously distributed in the fluid establishing a high accumulation zone at the base of the pipe; it can be observed that the solid particles gravitate at the bottom of the pipe when the inlet mean velocity is less, however when the mean velocity increases the solid particles become more buoyant in the fluid rather than settling down at the pipe bottom. Furthermore, the smaller particles are more homogenously distributed in the fluid than the larger particles because the larger particles tend to settle down at pipe bottom due to gravitational effect thereby forming a moving bed at lowest part of the pipe.
- Pressure drop increases with the increase in inlet mean velocity; because with the increase in mean velocity there in an increase in turbulence in the flow. Moreover, the pressure drop becomes higher when the solid volume fraction and particle size increases.
- From the S/N Ratio analysis of Taguchi's method optimal combinations of flow parameters have been achieved in order to optimize particle flow velocity and pressure drop individually.
- From ANOVA analysis it can be determined that for maximum particle flow velocity which is the desirable condition for smooth slurry flow and better productivity; the most influencing flow parameter is inlet mean velocity (contribution=98.38%)

followed by solid volume fraction (contribution=1.19%) and particle size (contribution=0.02%). For nominal pressure drop the most influencing flow parameter is found to be the solid volume fraction (contribution=48.41%) followed by inlet mean velocity (contribution=39.6%) and particle size (contribution=5.31%).

- A mathematical predictive model has been proposed for envisaging the responses analytically and it has been concluded that the proposed model can predict the responses with 6.95% error when compared to the experimental results.
- Genetic Algorithm has been coupled with the outcome of Taguchi's method in order to deal with the multi-objective optimization problem and a specific combination of flow parameters have been attained which can optimize the particle flow velocity and pressure drop simultaneously. The outcome of GA analysis has been compared with the simulated results and can be concluded that this strategy is efficient for optimizing multiple responses in a complex multiphase slurry flow problem. In addition it may be concluded that adopting this new GA approach particle flow velocity can be increased by 15% and pressure drop can be reduced by 29%.

The proposed mathematical model and outcomes of this current study can be applied for future academic research as well as in thermal power plants, food processing industries, chemical and pharmaceutical industries, slurry transporting industries and Oil and Gas industries in order to obtain an optimal pipe design. Further this new integrated Taguchi-GA approach would be helpful for optimizing pressure drop and particle flow velocity thereby initiating a smooth slurry flow which would reduce the power consumption/energy requirements as well as improve the transport efficiency in a complex multiphase slurry flow through horizontal pipelines.

ACKNOWLEDGEMENT

The authors would like to thank Amity University Haryana, India for providing support in carrying out this valuable research work.

REFERENCES

Chandel, S., Singh, S. N., & Seshadri, V. (2010). Transportation of high concentration coal ash slurries through pipelines. *International Archive of Applied Sciences and Technology*, *1*, 1-9.

Doron, P., Granica, D., &Barnea, D. (1987). Slurry flow in horizontal pipes-experimental and modeling. *International Journal of Multiphase Flow, 13*(4), 535-547.

Gillies, R. G., Shook, C. A., & Wilson, K. C. (1991). An improved two layer model for horizontal slurry pipeline flow. *The Canadian Journal of Chemical Engineering, 69*(1), 173-178.

Gillies, R. G., Shook, C. A., & Xu, J. (2004). Modelling heterogeneous slurry flows at high velocities. *The Canadian Journal of Chemical Engineering, 82*(5), 1060-1065.

Kaushal, D. R., & Tomita, Y. (2007). Experimental investigation for near-wall lift of coarser particles in slurry pipeline using γ-ray densitometer. *Powder Technology, 172*(3), 177-187.

Kaushal, D. R., Thinglas, T., Tomita, Y, Kuchii S., & Tsukamoto, H. (2012). CFD modeling for pipeline flow of fine particles at high concentration. *International Journal of Multiphase Flow, 43*, 85-100.

Kotcioglu, I., Cansiz, A., & Khalaji, M. N. (2013). Experimental investigation for optimization of design parameters in a rectangular duct with plate-fins heat exchanger by Taguchi method. *Applied Thermal Engineering*, *50*(1), 604-613.

Lahiri, K. S., & Ghanta, K. C. (2010). Slurry flow modelling by CFD. *Chemical Industry and Chemical Engineering Quarterly, 16*(4), 295-308.

O'Brien, M. P. (1933). Review of the theory of turbulent flow and its relations to sediment transport. *Transaction of the American Geophysical Union, 14*, 487-491.

Öktem, H., (2009). An integrated study of surface roughness for modelling and optimization of cutting parameters during end milling operation. *International Journal of Advance Manufacturing and Technology, 43*(9-10), 852-861.

Oroskar, A. R., & Turian, R. M. (1980). The critical velocity in pipeline flow of slurries. *American Institute of Chemical Engineers Journal*, *26*(4), 550-558.

Roco, M. C., & Shook, C. A. (1983). Modeling of slurry flow: the effect of particle size. *The Canadian Journal of Chemical Engineering*, *61*(4), 494-503.

Rouse, H. (1937). Modern conceptions of the mechanics of fluid turbulence. *American Society of Civil Engineers*, *102*, 463-505.

Toda, M., Ishikawa, T., Saito, S., & Maeda, S. (1973). On the particle velocities in solid-liquid two-phase flow through straight pipes and bends. *Journal of Chemical Engineering of Japan*, *6*(2), 140-146.

Turian, R. M., & Yuan, T. F. (1977). Flow of slurries in pipelines. *AIChE Journal, 23*(3), 232-243.

Wu, S. J., Ouyang, K., & Shiah, S. W. (2008). Robust design of microbubble drag reduction in a channel flow using the Taguchi method. *Ocean Engineering*, *35*(8-9), 856-863.

Pertanika J. Sci. & Technol. 27 (4): 2109 - 2121 (2019)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

An Eco-Innovative Green Design Method by QFD and TRIZ Tools- A Case Study of Brass-Ware Manufacturing

Anhad Malhotra, Sonu Rajak* and Sanjay Kumar Jha

Department of Production Engineering, Birla Institute of Technology, Mesra, Ranchi 835215, India

ABSTRACT

Eco-design is an advanced method for product and process to achieve sustainability. For sustainability, manufacturing companies should contemplate and encourage eco-design products and processes. Brass-ware manufacturing involved excessive energy consumption and material wastage, which makes the process unsustainable. The purpose of this study is to develop a systematic-innovation methodology for the eco-design process in brass-ware manufacturing by considering the eco-efficiency parameters based on Quality Function Deployment (QFD) approach and Theory of Inventive Problem Solving (TRIZ) tools. QFD is used for translating customers' requirements into manufacturing standards. TRIZ laws are used to solve the eco-contradiction matrix, which finds solutions to make the process eco-efficient and sustainable. On solving the eco-contradiction matrix generic solution is obtained such as '31 Porous Materials', '15 Dynamics', '10 Preliminary Action', and '40 Composite Material'. The application of the proposed method has been demonstrated by a case study of the brass-ware manufacturing process in India. This method may be useful and used as a potent tool for design engineers to invent new, robust, and eco-friendly products and processes.

Keywords: Eco-design, eco-innovation, quality function deployment, theory of inventive problem solving

ARTICLE INFO

Article history:
Received: 19 September 2018
Accepted: 21 June 2019
Published: 21 October 2019

E-mail addresses:
anhadmalhotra@gmail.com (Anhad Malhotra)
sonu.production@gmail.com, srajak.teqip@bitmesra.ac.in (Sonu Rajak)
sanjayjha@bitmesra.ac.in (Sanjay Kumar Jha)

ISSN: 0128-7680
e-ISSN: 2231-8526

INTRODUCTION

In today's world, environmental problems, namely global warming (Houghton et al., 2001), pollution, and the waste disposal problem (OECD, 2001) have become a serious issue. Post the industrial revolution and rapid strides in technology; manufacturers have not focused on the environment, thereby present systems become environmentally unsustainable. In fact, environmental issues have become critical for manufacturers. To make an environmentally friendly product and processes is becoming a major issue for the manufacturing companies. In the present scenario, it is required that every organisation/industry should be reducing environmental impacts caused by their products or process and also should be competitive to market; this activity is called eco-design. Hence, it has increased awareness among manufacturers for making environmental friendly process and products which would meet both customers demand and environmental requirements.

Systematic innovation for eco-design is an important step in the manufacturing organisation. To produce an eco-friendly product and process, the manufacturer should consider systematic innovation concept and environmental aspect in the early phase of design. Over the years the Indian brass-ware products have gained prominence throughout the world. A large number of individuals are employed in this sector. However, the methods and process involved in brass-ware manufacturing have not improved over the years. It is essential to develop a practical systematic-innovation method for eco-process design. In this context, the objective of this paper is to analyse and improve the process of brass-ware manufacturing, which should be eco-friendly and cost-effective. This study develops a systematic-innovation methodology for the green process by considering the eco-efficiency parameters based on QFD-I approach and the TRIZ tools.

Eco-friendly technologies are necessary to meet customer demand and environmental sustainability. The eco-friendly concept has been studied by many authors such as selection and design of sustainable transportation channels (Rajak et al., 2018) and design of circular production system (Vimal et al., 2019). Eco-innovative green materials have been considered by several authors (Ilyas et al., 2018a; Ilyas et al., 2018b; Ilyas et al., 2018c; Abral et al., 2019; Sanyang et al., 2018). A study carried out a detailed literature analysis on eco-design methods for developing new products based on QFD (Puglieri et al., 2011). In Ferrer et al. (2012), a technological eco-innovation preliminary design based on a computer-aided model for chemical engineering had been developed. Various methods have been considered to obtain the best results for eco-design (Kobayashi, 2006). Karlsson and Luttropp (2006) emphasized that eco-design produced an effective system solution, smart design of products, and methods. A recent study highlights environmental aspects into product/process development by using six sigma (Pusporini et al., 2013a). Later, a combination of six sigma and TRIZ has been presented by Wang et al. (2016) for the development of the new product. A study by Pusporini et al. (2013b) used green-QFD along with fuzzy logic to obtain an eco-friendly product. Environmental requirements and metrics have been used by Masui et al. (2013). Alemam and Li (2014) proposed an eco-design method by integrating QFD and Functional Analysis (FA) at the early stage of design. Chen and Liu (2003) presented an eco-innovative design method of products by considering the relationship between green-QFD, eco-efficiency, and TRIZ engineering parameters. The drawback of this paper is that they considered only for a product not for a process. Dehariya and Verma (2015) used green-QFD and House of Quality (HoQ) for designing an air-conditioner. A paper by Romli et al. (2015) presented an IEDM methodology formed from three stages namely, Life Cycle Assessment (LCA), eco-design process model, and enhanced eco-design QFD. A special eco-design HoQ is used to analyse the sustainability of manufacturing processes, product usage, and end-of-life strategy. Kobayashi (2005) introduced the software tool with a suitable parameter of the life cycle of the product to analyse the eco-design concept of the life cycle of the product. Alemam and Li (2016) dealt with the application of quality management systems for eco-design concepts. The paper used morphological charts, relation matrices as well as QFD for concept generation in eco-design. A study of the University of Taiwan developed an effective method for obtaining green products by integrating the green-QFD (Ko et al., 2016).

The 'Theory of Inventive Problem Solving' also known as TRIZ was developed by a Russian scientist Genrich Altshuller in the 1940s. It is a systematic thinking tool that can help to improve 'systems' in an innovative way. TRIZ is a potent tool to solve the inventive problems and has been used by many authors' over the past decade (Chen & Jiao, 2014; Pokhrel et al., 2015; Russo et al., 2015). A systematic literature survey based on TRIZ for transforming technical knowledge into a product concept has been conducted by Fiorineschi et al., (2018). TRIZ tool has been used by several authors such as the design of automotive parking brake lever (Mansor et al., 2014), for green supply chain problem (Moussa et al., 2017). A study discussed the guidelines needed to be taken into consideration for implementing the TRIZ laws of evolution for eco-design (Russo et al., 2011). Numerous design method and techniques, TRIZ tools, and eco-efficiency parameters to obtain desired results have been developed to support systematic innovation, which is summarised in Table 1.

Table 1
Different tools and techniques used for eco-design

Sl. No.	Tools	References
1	QFD	Kobayashi (2005); Puglieri et al. (2011); Pusporini et al. (2013a); Dehariya and Verma (2015); Romli et al. (2015); Alemam and Li (2016); Ko et al. (2016).
2	Morphological Charts	Alemam and Li (2016).

Table 1 *(Continued)*

Sl. No.	Tools	References
3	Matrix-Based Correlation	Alemam and Li (2014).
4	TRIZ Laws	Chen and Liu (2003); Russo et al. (2011).
5	Life Cycle Planning (LCP)	Kobayashi, (2005).
6	Life Cycle Assessment (LCA)	Kobayashi (2005); Dehariya and Verma (2015).
7	IEDM	Dehariya and Verma (2015)
8	Fuzzy Logic	Pusporini et al. (2013a)
9	Green- QFD	Chen and Liu (2003); Pusporini et al. (2013a); Alemam and Li (2014); Ko et al. (2016).
10	House of Quality (HoQ)	Pusporini et al. (2013a); Ko et al. (2016).

Based on the literature review, to the best knowledge of the authors, it has been found that researchers have successfully integrated QFD and TRIZ Laws for eco-design of products. However, no significant methodology has been developed for systematic and innovative eco-design of a process. Moreover, no concrete research has been carried out to obtain the green process for brass-ware manufacturing. This paper aims at the integration of Green-QFD, TRIZ tools, and eco-efficiency parameters for obtaining a green process for brass-ware manufacturing. Further, the rest of the paper has been organized as follows: The methodology has been described in Section 3. In Section 4, a case study of the brass-ware manufacturing process is introduced. Results and discussion are shown in Section 5. Finally, Section 6 concludes the paper with future works.

METHODOLOGY

In order to develop an eco-friendly process, a systematic-innovation eco-design concept was used by integrating QFD-I and TRIZ tools. The QFD-I method can be used to systematically design the problem, analyse the technical feasibility, and validate the design solution by using the House of Quality (HoQ). TRIZ is a systematic method for creating innovative resolutions. In this paper, the blend of QFD-I and TRIZ method has been used to construct a framework for the new eco-design process. QFD-I has been adopted to analyse the manufacturer's need and transfer the problem into requirements. TRIZ is used to establish an eco-contradiction matrix, which can be used to identifying the conflict between customers' need and eco-efficiency elements for deriving TRIZ invention principles.

Systematic-Innovation Design Process

Product design engineers must consider both technical innovation and environmental consideration while designing innovative product and processes. Seven principles have been identified by the World Business Council of Sustainable Development (WBCSD) to develop eco-friendly product or processes to reduce the environmental impact, which is as follows:

(i) Material reduction- a reduction of the material intensity of its product, processes, and services.

(ii) Energy reduction- a reduction of energy intensity of its product, processes, and services.

(iii) Toxicity reduction- reduction in the dispersion of any toxic materials.

(iv) Material retrieval- enhancing the recyclability of materials.

(v) Resource sustainable- maximize the sustainable use of renewable resources.

(vi) Product durability- extend the durability of its products, processes, and services.

(vii) Product service- Increasing the service intensity of its product, processes, and services.

The proposed method had been divided into three phases. In the first phase, the seven eco-efficiency elements for the design of a product or process were defined and assigned respective TRIZ numbers. In the second phase, customers' need and quality characteristics had been identified using the QFD-I method. In the third phase, the standard solution of the problem had been obtained by using the eco-contradiction matrix and TRIZ inventive principles. The proposed method might be helpful for the design engineer to produce novel solution idea for the eco-design of product and processes.

QFD-I Approach

Any organisation aims to produce quality products and services. In today's competitive environment, quality is a requirement that customers expect. QFD is a critical aspect of the quality control policy of an organisation. QFD is a very powerful tool as it incorporates customer needs into the design parameters so the final product will be better designed to meet customer expectations. The main concept of traditional QFD-I is to consider four relationship matrices that included product planning, part planning, process planning, and production planning matrices, respectively. Each translation uses a matrix, also called the House of Quality (HoQ). In this paper, the product planning matrix was constructed to obtain the desired quality characteristics.

Eco- Contradiction Matrix

The steps of the eco-contradiction matrix are shown as follows:

STEP 1: Quality characteristics are placed in the improving column and the eco-efficiency elements into the worsening row.

STEP 2: Quality characteristics and eco-efficiency elements are replaced with standard features in TRIZ intuitively.

STEP 3: Set contradiction marks (X) between the technical terms in TRIZ.
STEP 4: Derive invention principles based on the TRIZ contradiction matrix.
STEP 5: Develop eco-innovation product concept based on the invention principles.

CASE STUDY

Problem Environment: Application of the new Method on Brass-Ware Manufacturing Process

Brass-ware product manufacturing has been prevalent in India since times immemorial. Brass–ware manufacturing has many key parameters that affect the end product. These include energy, labour, processes employed, and machinability. In addition to these, the economical sourcing of raw materials as well as minimize environmental effects plays a vital role in brass–ware manufacturing. Therefore, the manufacturing of brass-ware incorporates all areas which would have a bearing in green manufacturing.

The problem environment was studied to convert a traditional brass-ware manufacturing process to a green process by applying QFD-I coupled with the TRIZ laws. The environmental concern was to identify the key parameters such as product volume, labour requirement, and reusability that would help in making the brass-ware manufacturing process sustainable without impacting the quality and uniqueness of the products.

Steps to carry out the analysis were as follows: -

- House of Quality (HoQ): - This step involved identifying what's (Environmental Requirements) and how's (Engineering Metrics) that would affect the brass–ware manufacturing. The what's are the parameters to make the process green and the How's are the factors that can be changed by the manufacturer to make the process green. Both the what's and how's were selected as presented by Masui et al. (2013) for environmental requirements and engineering metrics. Once the metrics were identified, the 9–3–1 weighting method was used to signify the correlation between the what's and how's with respect to making the process green. For example, the 'Weight of the Product' was strongly correlated with 'Less Material Usage,' and hence the weight of '9' had been assigned. These weights were obtained by gaining insights about the brass- ware manufacturing process using literature review as well as in discussion with industry experts and decision makers. Using the weight, the raw score was calculated, which was done by multiplying the weights of each what's with the weights of the respective how's and adding the same. The raw score was indicative of the importance of the quality characteristics of making the process green. For better understanding, a relative score was calculated. The higher the relative score, the more important the engineering metric and that makes the process green. QFDE for the brass-ware manufacturing process is given in Table 2. The relative score has been computed as:

$$\text{Relative Score} = {}^{\text{Raw Score}}\!/_{\sum \text{of Raw Scores}} \tag{1}$$

Table 2

QFDE for brass–ware manufacturing process

		How's (Engineering Metrics/Quality Characteristics)								
What's (Environmental Requirement)	Importance to Manufacturer (Weights)	The weight of the product	The volume of the product	Number of Parts	Number of Types of materials	Hardness	Physical Lifetime	Amount of energy	Rate of Recycled	Biodegradability
Less Material Usage	9	9	9	9						3
Easy to transport and retain	9	9	9	9	3		3		1	
Easy to process and assemble	9	3		9	3		1	9	3	
Less energy consumption	9				3	3	9	9		
Ease to disassemble	3	9	9	9	3				3	
Easy to clean	3									
Less labour	9	3	3	9				3		
Easy to reuse	3								9	
Safe to incinerate	1									
Safe to landfill/dispose easily	9	9			3					9
Harmless/safe emission	3								3	
Raw score		324	216	351	90	27	117	189	72	81
Relative score		0.22	0.15	0.24	0.06	0.018	0.08	0.13	0.05	0.06

Using the Relative Scores, the following quality characteristics were selected:

- Product weight
- Number of parts
- Product volume
- Amount of energy consumption
- Physical lifetime
- Eco-Contradiction Matrix: - The top five relative scores obtained were then used to make an eco–contradiction matrix. Each of the five-engineering metrics obtained from the HoQ were given respective TRIZ number. Similarly, the seven eco-efficiency parameters were assigned respective TRIZ number. Thus, the contradiction matrix was formed and solved using the 40 inventive principles. Explanation of each contradiction is given as follows:
    - X1: - 'Product Weight' is strongly correlated with 'Material Reduction,' which means that if the product weight increases, the quantity of material used per unit service increases, that makes the product less eco-friendly.

● X2: - Increasing the number of parts would lead to increase in product durability. If the number of parts increases, the reliability of the product may decrease, leading to the failure of the product.

● X3: - 'Product Volume' is strongly correlated with 'Material Retrieval,' which means that as the volume of the product increases, the amount of material that can be recycled will be less, making the product less eco-friendly.

● X4: - As Energy Consumption Increases, the product becomes less eco-friendly.

● X5: - Increasing the lifetime of the product has a direct correlation with the 'Product Service.'

All the above correlations had been solved using an eco-contradiction matrix. Each contradiction had been solved using the 40 inventive principles. The TRIZ eco-contradiction matrix is shown in Table 3.

Table 3
TRIZ eco-contradiction matrix

Eco-Contradiction Matrix				**Worsening Parameters** Eco-Efficiency Elements						
				A. Material Reduction	B. Energy Reduction	C. Toxicity Reduction	D. Material Retrieval	E. Sustainable Resource	F. Product Durability	G. Product Service
		Item	TRIZ No.	23	20	31	2	34	27	33
Improving Parameters	Quality Characteristics	Product Weight	1	X1						
		Number of Parts	4						X2	
		Product Volume	8				X3			
		Amount of Energy Consumption	20		X4					
		Physical Lifetime	27							X5

RESULTS AND DISCUSSIONS

Table 4 was obtained on solving the eco-contradiction matrix. Each solution of the five contradictions had been explained.

Table 4
Results of the eco-contradiction matrix

Contradiction	Invention Principle	Generic Solution
X1= (1,23)	(3,5,31,35)	31. Porous Materials
X2= (4,27)	(15,28,29)	15.Dynamics
X3= (8,2)	(10,14,19,35)	10.Preliminary Action
X4= (20,20)	-	-
X5= (27,33)	(17,27,40)	40.Composite Material

- X1 is the contradiction of "increasing the weight of the product without affecting the material reduction." Four main solutions are available to solve this contradiction:

1. Local quality (3)
2. Merging (5)
3. Porous materials (31)
4. Parameter changes (35)

On analysing each solution, the most suited is **"porous material."**

"Porous Material "can be explained as:

Make an object porous or add porous elements (inserts, coatings, etc.). Example: - Drill holes in the structure to reduce weight

Additionally, if the product is porous, one can use the pores for some useful function.

In the brass-ware manufacturing process for statues, use of investment casting instead of sand casting is an anticipated solution for making the process green. It helps in obtaining a smooth finish and reducing the weight of the product (Hollow).

- X2 is the contradiction of "increasing the number of parts without affecting the product durability."

Three main solutions are available to solve this contradiction:

1. Dynamics (15)
2. Mechanics substitution (28)
3. Pneumatics and hydraulics. (29)

On analysing each solution, the most suited is **"dynamics."**

Dynamics deals with designing the characteristics of the process in such a way that if the process is rigid or inflexible, it will make it movable or adaptive. It implies that as the parts increase, each part would be made movable/adaptive, ensuring that durability is not affected.

- X3 is the contradiction of "increasing the volume of the product without affecting the material retrieval." Four main solutions are available to solve this contradiction:

1. Preliminary action
2. Spheroidality
3. Periodic action
4. Parameter changes.

On analysing each solution, the most suited is **"preliminary action."**

Preliminary action can be explained as:

1. Perform, before it is needed, the required change of an object (either partially or fully). Example: - Sterilize all instruments needed for a surgical procedure on a sealed tray.
2. Pre-arrange objects such that they can come into action from the most convenient place and without losing time for their delivery.
   Example: - Kanban arrangements in a Just-in-Time (JIT) factory.

In the brass-ware manufacturing process, **"design for manufacturability"** is an anticipated solution for making the process green. It will help in:

1. Optimising all the manufacturing functions such as fabrication, assembly, test, procurement, shipping, delivery, service, and repair, and
2. Assure the best cost, quality, reliability, regulatory compliance, safety, time-to-market, and customer satisfaction

- X5 is the contradiction of "increasing product lifetime without affecting product service."

Three main solutions are available to solve this contradiction:

1. Another dimension
2. Cheap short living
3. Composite material.

The most suitable solution is the use of "**composite material"** to obtain the desired product lifetime. In brass-ware manufacturing use of corrosion-resistant brass for harsh environments is an anticipated solution for making the process green.

Many systematic innovation methods have been discussed in the literature, but these methods are only focused on eco-design of a product (Chen and Liu, 2003; Ko et al., 2016). The eco-friendly process is also an important aspect need to be explored by the design engineer in the starting phase of design, such as planning and conceptual design phases. Based on the literature survey, to the best knowledge of the authors, it has been found that no concrete research has been carried out for the eco-friendly process. In this context, this paper presents the eco-innovative green design method for brass-ware manufacturing.

CONCLUSIONS

In the pursuit of sustainable and green manufacturing, there is a need to examine the eco-design of both products and processes. The contradiction is the biggest impediment to a new generation of solutions. The design process integrating quality characteristics of QFD-I and eco-efficiency elements into the proposed eco-contradiction matrix has been established for solving eco-innovative design problem. In this paper QFD-I, TRIZ tools, and eco-efficiency parameters were effectively used to make brass-ware manufacturing a green process. After, solving the eco-contradiction matrix, it has been found that by considering the solution such as the use of porous material, dynamics, preliminary action, and composite material, brass-ware manufacturing process may be made green. A case study of brass-ware manufacturing demonstrated the applicability and feasibility of the systematic-innovation method. Further, the approach could be applied to any type of other manufacturing processes.

FUTURE WORK AND RECOMMENDATION

Future work for this method has the potential to be adopted for numerous applications. Various tedious processes can be easily analysed and made environment-friendly. Fuzzy logic, as well as LCA, can be applied to get better results. Besides the considered parameters, other parameters can also be taken into consideration, which would enhance the result.

ACKNWOLEDGEMENT

Authors would like to express their thanks to Brass-ware manufacturing company for allowing to carry out this research work.They would also like to offer their sincere thanks to all Professors/teaching staff and non-teaching staff of the Department of Production Engineering, Birla Institute of Technology, Mesra, India for their support and encouragement.

REFERENCES

Abral, H., Basri, A., Muhammad, F., Fernando, Y., Hafizulhaq, F., Mahardika, M., ... & Stephane, I. (2019). A simple method for improving the properties of the sago starch films prepared by using ultrasonication treatment. *Food Hydrocolloids, 93*, 276-283.

Alemam, A., & Li, S. (2014). Integration of quality function deployment and functional analysis for eco-design. *International Journal of Mechanical Engineering and Mechatronics, 2*(1), 1-12.

Alemam, A., & Li, S. (2016). Matrix-based quality tools for concept generation in eco-design. *Concurrent Engineering, 24*(2), 113-128.

Chen, J. L., & Jiao, W. S. (2014). TRIZ innovative design method for eco-leasing type product service systems. *Procedia CIRP, 15,* 391-394.

Chen, J. L., & Liu, C. C. (2003, August 19-21). An eco-innovative design method by green QFD AND TRIZ tools. In DS 31: Proceedings of ICED 03, the *14th International Conference on Engineering Design* (pp.1-10). Stockholm, Sweden.

Dehariya, P. K., & Verma, D. S. (2015). An application of green quality function deployment to designing an air conditioner. *International Journal of Engineering Research and Applications, 5*(7), 147-152.

Ferrer, J. B., Negny, S., Robles, G. C., & Le Lann, J. M. (2012). Eco-innovative design method for process engineering. *Computers and Chemical Engineering, 45,* 137-151.

Fiorineschi, L., Frillici, F. S., & Rotini, F. (2018). Enhancing functional decomposition and morphology with TRIZ: Literature review. *Computers in Industry, 94,* 1-15.

Houghton, J. T., Ding, Y. D. J. G., Griggs, D. J., Noguer, M., van der Linden, P. J., Dai, X., ... & Johnson, C. A. (2001). *Climate Change 2001: The Scientific Basis.* Cambridge, UK: Cambridge University Press.

Ilyas, R. A., Sapuan, S. M., & Ishak, M. R. (2018). Isolation and characterization of nanocrystalline cellulose from sugar palm fibres (Arenga pinnata). *Carbohydrate Polymers, 181,* 1038-1051.

Ilyas, R. A., Sapuan, S. M., Ishak, M. R., & Zainudin, E. S. (2018). Development and characterization of sugar palm nanocrystalline cellulose reinforced sugar palm starch bionanocomposites. *Carbohydrate Polymers, 202,* 186-202.

Ilyas, R. A., Sapuan, S. M., Ishak, M. R., & Zainudin, E. S. (2018). Sugar palm nanocrystalline cellulose reinforced sugar palm starch composite: Degradation and water-barrier properties. *IOP Conference Series: Materials Science and Engineering, 368*(1), 012006.

Karlsson, R., & Luttropp, C. (2006). EcoDesign: what's happening? An overview of the subject area of EcoDesign and of the papers in this special issue. *Journal of Cleaner Production, 14*(15-16), 1291-1298.

Ko, Y. T., Chen, M. S., & Lu, C. C. (2016). A systematic-innovation design approach for green product. *International Journal of Construction Engineering and Management, 5*(4), 102-107.

Kobayashi, H. (2005). Strategic evolution of eco-products: a product life cycle planning methodology. *Research in Engineering Design, 16*(1-2), 1-16.

Kobayashi, H. (2006). A systematic approach to eco-innovative product design based on life cycle planning. *Advanced Engineering Informatics, 20*(2), 113-125.

Mansor, M. R., Sapuan, S. M., Zainudin, E. S., Nuraini, A. A., & Hambali, A. (2014). Conceptual design of kenaf fiber polymer composite automotive parking brake lever using integrated TRIZ–Morphological Chart–Analytic Hierarchy process method. *Materials and Design (1980-2015), 54,* 473-482.

Masui, K., Sakao, T., Kobayashi, M., & Inaba, A. (2003). Applying quality function deployment to environmentally conscious design. *International Journal of Quality and Reliability Management, 20(1),* 90-106.

Moussa, F. Z. B., Rasovska, I., Dubois, S., De Guio, R., & Benmoussa, R. (2017). Reviewing the use of the theory of inventive problem solving (TRIZ) in green supply chain problems. *Journal of Cleaner Production, 142,* 2677-2692.

OECD. (2001). *OECD Environmental Outlook.* Organisation for Economic Co-operation and Development, Paris, France.

Pokhrel, C., Cruz, C., Ramirez, Y., & Kraslawski, A. (2015). Adaptation of TRIZ contradiction matrix for solving problems in process engineering. *Chemical Engineering Research and Design, 103,* 3-10.

Puglieri, F. N., Ometto, A., & Miguel, P. C. (2011). Eco-design methods for developing new products based on QFD: a literature analysis. *Product: Management and Development, 9*(1), 23-29.

Pusporini, P., Abhary, K., & Luong, L. (2013a). Development of Environmental Performance Model Using Design for Six Sigma (DFFS). *International Journal of Materials, Mechanics and Manufacturing, 1(1),* 102-106.

Pusporini, P., Abhary, K., & Luong, L. (2013b). Integrating environmental requirements into quality function deployment for designing eco-friendly product. *International Journal of Materials, Mechanics and Manufacturing, 1*(1), 80-84.

Rajak, S., Parthiban, P., & Dhanalakshmi, R. (2018). Selection of Transportation Channels in Closed-Loop Supply Chain Using Meta-Heuristic Algorithm. *International Journal of Information Systems and Supply Chain Management (IJISSCM), 11*(3), 64-86.

Romli, A., Prickett, P., Setchi, R., & Soe, S. (2015). Integrated eco-design decision-making for sustainable product development. *International Journal of Production Research, 53*(2), 549-571.

Russo, D., Regazzoni, D., & Montecchi, T. (2011). Eco-design with TRIZ laws of evolution. *Procedia Engineering, 9,* 311-322.

Russo, D., Schöfer, M., & Bersano, G. (2015). Supporting ECO-innovation in SMEs by TRIZ Eco-guidelines. *Procedia Engineering, 131,* 831-839.

Sanyang, M. L., Ilyas, R. A., Sapuan, S. M., & Jumaidin, R. (2018). Sugar palm starch-based composites for packaging applications. In M. Jawaid & S. K. Swain (Eds.), *Bionanocomposites for packaging applications* (pp. 125-147). Cham, Switzerland: Springer International Publishing AG.

Vimal, K. E. K., Rajak, S., & Kandasamy, J. (2019). Analysis of network design for a circular production system using multi-objective mixed integer linear programming model. *Journal of Manufacturing Technology Management, 30(3),* 628-646.

Wang, F. K., Yeh, C. T., & Chu, T. P. (2016). Using the design for Six Sigma approach with TRIZ for new product development. *Computers and Industrial Engineering, 98,* 522-530.

Pertanika J. Sci. & Technol. 27 (4): 2123 - 2138 (2019)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Utilisation of *Cassia surattensis* Seeds as Natural Adsorbent for Oil Content Removal in Oilfield Produced Water

Mohmad Khalel Ibrahim[1], Ali Abd Al-Hassan[2] and Ahmed Samir Naje[3]*

[1]*Department Plant Protection, College of Agriculture, University of Al Muthanna, Samawah, Muthanna 66001, Iraq*
[2]*Department of Chemical Engineering, College of Engineering, University of Al Muthanna, Samawah, Muthanna 66001, Iraq*
[3]*Collage of Water Resource Engineering, AL-Qasim Green University, Babylon 51031, Iraq*

ABSTRACT

Oil content removal by *Cassia surattensis* seeds was deliberated in a batch. Significance of the current study is to treat water that is associated with oil. The water contains high organic materials, which is completely dissolved. Previous processors had problems such as low efficiency and costly. *Cassia surattensis* seed was branded by using Fourier transform infrared spectroscopy. Contact time, dose quantity and pH value were examined for oil content elimination in produced water (PW). A declaration of pH value from 2.0–9.0 was accompanied by an improved amount in oil that was adsorbed. The high oil content removal was 79.4% at pH 2, 120 min and 2 g of *Cassia surattensis* seeds. The equilibrium adsorption values were obtained and studied by using the Freundlich and Langmuir models. The Freundlich model was found to obtain the finest relation with the tested values. The greatest adsorption capacity of the oil content was 12.18 mg/g. The outcome of the statistical model showed successful results, whereby the F-value was 51.9 and p-value was 0. The present study results proposed that *Cassia surattensis* seeds may be used as a low-cost adsorbent to eliminate oil content in produced water.

Keyword: Adsorption, natural adsorbent, produced water, water treatment

ARTICLE INFO

Article history:
Received: 11 May 2019
Accepted: 03 July 2019
Published: 21 October 2019

E-mail addresses:
alkhafajia3@yahoo.com (Mohmad Khalel Ibrahim)
ali_salih_iraq@yahoo.com (Ali Abd Al-Hassan)
ahmednamesamir@yahoo.com (Ahmed Samir Naje)
* Corresponding author

INTRODUCTION

Water contamination is a unique and one of the world's most unwanted ecological problems that needs solution (Aljeboree et al., 2017). Produced water (PW) is a type of

ISSN: 0128-7680
e-ISSN: 2231-8526

wastewater that has a major effect on the environment and oil production (Hosny et al., 2016), The PW properties vary significantly depending on the sort of organic creation that is being shaped, geological creation as of where the water was bent and the site (Mousa et al., 2017). PW comprises many organic, inorganic and heavy metal components. This water discharge can contaminate the surface, subversive water and soil (Fakhru'l-Razi et al., 2009; Shokrollahzadeh et al., 2012). Oil production produced great amounts of wastewater (Hassan et al., 2018). The approximate volume of PW may surpass the volume of crude oil produced by extra of two eras than throughout the lifetime of the reservoir (Bagheri et al., 2018). PW management is self-same significant, owing to lawmaking and ecological anxieties. All along additional strict environmental rules on diverse PW treatments need to be released earlier from the oil and gas manufactures, including previous prevention in reservoirs to decrease damage (Fathy et al., 2018).

The treatment goals are to recover the dissolved oil and edge the last waste for reuse or discard (da Silva et al., 2015) . The wastewater from oil productions are gutted by many physical and chemical incomes before the release and rules place severe bounds on stages of pollutants which can be cleared to the sea (Bakke et al., 2013).Conformist treatment skills were industrialised to eliminate the organic content from water and PW, including a decrease followed by chemical oxidation (Naeem et al., 2018), flotation /sedimentation (Jiménez et al., 2017), electrocoagulation (Hernández-Francisco et al., 2017), bacterium (Li et al., 2005) and adsorption/ membrane (Kusworo et al., 2018). Amongst the novel and emerging know-hows that are available is adsorption by agricultural leftover products. Adsorption is considered as a cost active and well-organised choice. Meanwhile the adsorbent materials must obviously be obtainable and inexpensive (Dehghani et al., 2016).

Commercially activated carbon has been studied as an adsorbent for the removal of organic content from wastewater for several years due to the great specific surface area and pore structure, but it is quite expensive. The discovery of an alternative adsorbent to replace the costly activated carbon is highly encouraged. Nowadays, research are focusing more on various adsorbents which have organic-binding capacities and able to remove the oil content from PW at a lower cost. The production of lignin obtained from industrial waste fulfills the requirement of a low-cost adsorbent (El Messaoudi et al., 2016).

The use of non-conventional, low-cost adsorbents prepared from agricultural wilds and spin-offs can-not alone decrease a big amount of solid waste, nonetheless to be actual attractive. Certain squat-cost adsorbents comprise sugarcane bagasse (Pereira et al., 2010) , tea waste (Pirbazari et al., 2014), orange peel (Feng et al., 2009a; Feng et al., 2009b) , longan seed (Yang et al., 2015) , jackfruit peel (Foo & Hameed, 2012), biochars (Inyang et al., 2012), groundnut hull (Owalude & Tella, 2016) , potato peels (Kyzas et al., 2016), corn silk (Yu et al., 2016) and *Cassia surattensis* flowers (Maurya et al., 2018). The types of Cassia include about 600 classes and are extensively dispersed. They are universally recognized as health and have various biological and pharmacological possessions. *Cassia surattensis* is categorised in the Fabaceae family. These flowering florae are extensively grown as ornamental plants in hot and subtropical zones. This plant is usually on rummage-sale in many republics for food and medicinal use. No native uses were recognised for *Cassia surattensis*, but the bay and leaves are said to be antiblenorrhagic. It is also a stuff for a decoction of the origins reference (Kumar et al., 2014). The objective of the present study is to illuminate the elimination of oil content from PW through the adsorption on *Cassia surattensis* seeds. The consequence of some constrains, for instance, cassia dose, pH value and contact time was measured. Furthermore, the adsorption mechanism on bore size is explained by using the Langmuir and Freundlich isotherms.

MATERIALS AND METHODS

Chemicals and Reagents

All the substances and reagents rummage-sale were of a rational score, NaOH (98% purity) and H_2SO_4 (98% purity) were bought from India. *Cassia surattensis* seeds were obtained from the University Muthanna in Iraq and were identified at the laboratory of chemical disciplines. Pods were dehydrated and kernels were separated from the shells. The seeds were washed with distilled water to remove dirt prior to ventilation. Dried *Cassia surattensis* seeds were ground and dried in the sun for 48 h and used as the new adsorbent. The sun dried seeds of the adsorbent was preserved with HCL concentration of 0.1 M for 4 hand was systematically washed with distilled water until it obtained a neutral pH. Then the adsorbent seeds were washed with distilled water. *Cassia surattensis* seeds as the approved adsorbent were sieved with a 0.3 mm mesh sieve (Besmak sieves from 2.36mm to 0.075mm).

Fourier Transform Infrared (FTIR) Analysis

Fourier transform infrared (FTIR) analysis was applied to determine the surface functional groups, by using FTIR spectroscope (FTIR-2000, Bruker), whereby the spectra were recorded from 4000 cm^{-1} to 500 cm^{-1}.

Produced Water (PW)

The produced water (PW) was taken from the Al-Ahdab oilfield, in Iraq. The PW was strained to remove most of its solids and it was formerly set aside at 5ºC to ensure that the physiognomies for PW determination were not tapered or worn. PW description is expected in Table 1.

Table 1
Characteristics of produced water

Parameter	Value
Oil	118 (ppm)
Turbidity	102.4 NTU
pH	7.11
Solution oxygen content	0.051 (ppm)
Specific gravity	0.996
conductivity	60688.64 µs/cm
TSS	18.2 (ppm)
TDS	38840.73 (ppm)
Viscosity	1.101 m Pa/S
Iron	0.31 (ppm)
Sulphate	58.4 (ppm)
Manganese	2.5 ppm
Chrome	0.14 ppm

Experimental Procedure

The effects of experimental parameters, such as adsorbent dose (0.5-3 gm.), pH value (2.0 – 9.0) and contact time (30 – 150 min) on the adsorptive oil content removal was studied in a batch. All adsorption experiments were conducted in 250 mL conical flasks, with added 150 mL of produced water solution (at the desired concentration and pH value) . Water from the oilfield was shaken on a magnetic shaker at (25ºC ± 3ºC) and stirred at 200 rpm. Wastewater was engaged at the required intermissions and a centrifuge device was used to unravel the wastewater in two distinct layers. The effect of the solution pH value on the oil adsorption was examined in a similar way. However, the initial pH value of the solutions was in accordance with the varying pH values of 2.0 – 9.0 with the addition of either 0.1 M NaOH or 0.1 M H_2SO_4. The pH value standards were also logged when the *Cassia surattensis* seeds -oil postponements had touched equilibrium. Diverse amounts of *Cassia surattensis* seeds were in the range of 0.5 g – 3 g more than the PW. The amount of PW at symmetry q_e(mg/g) was anticipated from the mass balance, which was expected as from below (Naeem & Hassan, 2018):

$$q_e = \frac{V(C_\circ - C_e)}{W} \tag{1}$$

Where, q_e(mg/g) is all oil in produced water (PW) per mass unit of watermelon adsorbent at convinced time (t), V (ml) is the wastewater volume , W is the weight of *Cassia surattensis* seeds (g) and $C_\circ$ and C_e (ppm) are the unique and at time t concentration of wastewater congruently. The oil abolition by *Cassia surattensis* seeds was anticipated at all equilibration through the following equation:

$$Adsorption\ (\%) = \frac{C_{\circ} - C_e}{C_{\circ}} \ \text{x}\ 100 \qquad (2)$$

Logical Measurements

The oil content in PW was determined through a UV–spectra meter (UV-1800 Shimadzu, Japan) that was connected to a computer at all-out absorption wavelength (218 nm) for n-hexane .The turbidity was slow-moving and measured by using a turbid meter (Lovibond, SN 10/1471, and Germany), which would read the turbidity. The pH measure was mainly by using a pH meter (Model 2906, Jenway L_{td}, UK).

Organic Testing by using a UV-spectra Meter

PW of 50 mL was placed in a closed cylinder to avoid the oil emulsion from disturbance. Then 5 ml of n-hexane of below the acidic condition (pH2) was added,followed by an energetic shake for 2 min. After 10 min, the solution was separated from the two distinct layers, in which the above layer (oil) was engaged for the absorbance dimension. Later from the calibration curve the oil content in PW was determined.

Statistical Method

Statistical technique method, response surface methodology (RSM) as well as a statistical program (Minitab-17) were used to design the experiment and predict the operational factor influence individually and interactively. The main impacts of these factors, such as adsorbent dose (X_1), adsorption time (X_2) and pH (X_3), were studied according to their ranges, as shown in Table 2. A central composite design (CCD) method was employed to create the second order response surface correlation which was related to the operational parameters (Equation 5) as follows, whereby the indicator of the model quality is the coefficient of variance (R^2):

$$Y = B_0 + \sum_{i=1}^{q} B_i X_i + \sum_{i=1}^{q} B_{\boldsymbol{i}} X_i^2 + \sum_{i} \sum_{j} B_{\boldsymbol{j}} X_i X_j + \varepsilon \qquad (3)$$

Where, Y is the studied responses; X_1, X_2, to X_q are the operational variables; B_o is a regression constant, B_i is the linear regression coefficient, B_{ii} is the squared regression coefficient and B_{ij} is the cross-product regression coefficient; **ε** is a random error.

Table 2
Operational parameters

Parameters	Ranges
X_1: Adorbent dose (gm)	0.5-3
X_2: adsorption time (min)	30-150
$X_{3\,:}$ pH	9-2

Adsorption Isotherms

The discovery of suitable relations was aimed at the equilibrium facts, which is important to improve the adsorption scheme design for oil elimination. The models of adsorption were mentioned in the relation between the adsorbate concentration in the wastewater and the quantity of adsorbate which was adsorbed by the unit mass of adsorbent at a constant temperature. The adsorption isotherm model labels the adsorption procedure through information on the adsorption mechanisms, surface possessions and adsorbent empathies. The original adsorbents were industrialised in the current study. Therefore it was fundamentally required to assess the equilibrium data with the diverse significant isotherm models. In the present work two limited models were Langmuir and Freundlich were used (Rangabhashiyam & Selvaraju, 2015; Bediako et al., 2015).

Langmuir Model. The Langmuir model accepts one contaminant particle determination. It alone inhabits one lively place on the homogeneous adsorbent surface (Dahri et al., 2015) . The Langmuir model equation is as follows:

$$q_e = \frac{q_m K_L C_e}{1 + K_L C_e} \tag{4}$$

Where, C_e is the equilibrium attentiveness (ppm) and q_e is the amount adsorbed per specified amount of adsorbent (mg/g), K_L is the Langmuir equilibrium constant.

Freundlich Model. The Freundlich isotherm shoulders that the adsorption occurs on heterogeneous surfaces (Naeem & Hassan, 2018). The Freundlich model calculation is as follows:

$$q_e = K_F C_e^{1/n} \tag{5}$$

Where, K_F is Freundlich equilibrium constant, n is an empirical constant and disruption of the footings have the usual sense**.**

RESULTS AND DISCUSSION

FTIR Analysis

Cassia surattensis seeds were aimed to examine the functional group with FTIR. The peak sizes in the spectrum were direct indications of the quantity of multiple presences. The consequences of using FTIR established that there were many typical functional groups in the *Cassia surattensis* seeds. Figure 1 exposes the complete band at 3363.67 cm^{-1} in *Cassia surattensis* seeds which was accredited due to amine ($-NH_2$) or hydroxyl ($-OH$) stretching or flared of polymeric mixes. This band appeared in the mediocre area at 2933.83 cm^{-1} –2904.89 cm^{-1} in the FTIR spectra of *Cassia surattensis* seeds was accredited to the presence of the $C-H$ bond. The $C-O$ flared of *Cassia surattensis* seeds through the adsorption technique at 1732.13 cm^{-1}. The $C-O$ widening was detached to a higher occurrence in place of a consequence of carboxyl ($^{-}C-O$) group participation in the adsorption technique of oil content with *Cassia surattensis* seeds. 1510.13 cm^{-1} designated the aromatic rings, whereas 1425.44 cm^{-1} and 1373.43 cm^{-1} were connected by the $C-O$ in phenols and $-CH_3$, congruently. Formerly the modern bands at below 800 cm^{-1} were finger print regions of phosphate and sulphur functional groups (Jafer et al., 2019).

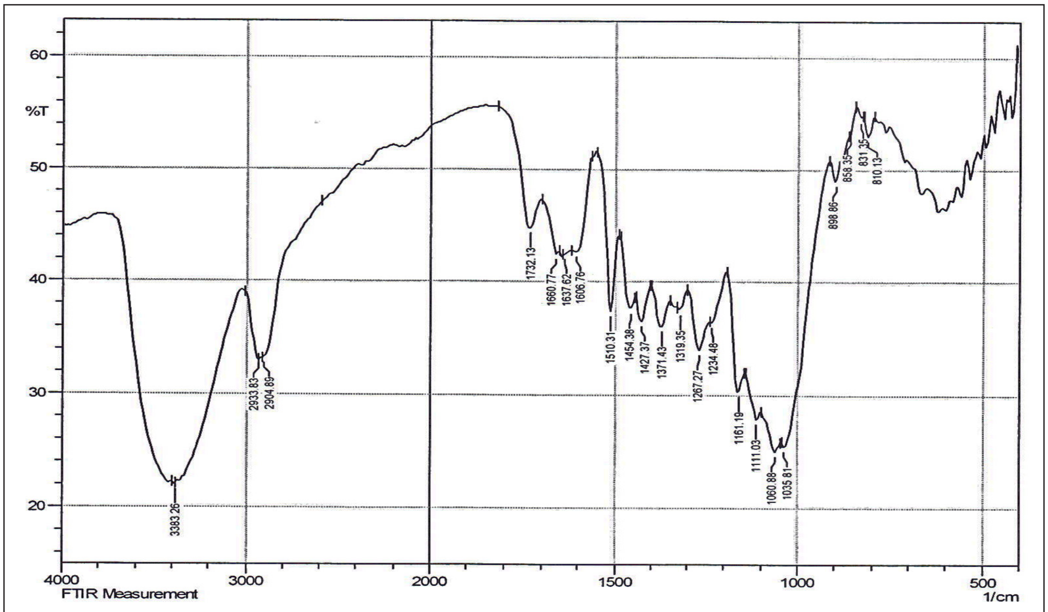


Figure 1. FTIR investigation of cassia surattensis seeds before adsorption for 0.3 mm size

Modeling of Degradation Removal

The effects of different issues on organic adsorbent by adsorption processes were investigated by using the combined conditions. Twenty experimentations were completed by using the response surface method to design the batch experiments as shown in Table

3. The equation which predicted the organic concentration was attained according to the response surface experimental design, applied relation between the independent variables and response, based on the coded unit in the experimental results, was given by a third-order model as follows:

$$Oil\_removal = 46.1 + 30.65X_1 + 0.177X_2 - 6.31X_3 - 8.28X_1^2 - 0.000010X_2^2 + 0.487X_3^3 + 0.0032X_1X_2 + 0.009X_1X_3 - 0.02637X_2X_3 \quad (6)$$

The values of the operational variables, percentage removal of the studied responses, i.e. oil removal, dose, adsorption time and pH are as shown in Table 3.

Table 3
Results of the studied variables

Run	X_1: Adsorbent dose (gm)	X_2: Adsorption time (min)	X_3 (pH)	oil removal (%)
1	0.5	30	2	52
2	3	30	2	57
3	0.5	150	2	68
4	3	150	2	72.6
5	0.5	30	9	38.9
6	3	30	9	42.7
7	0.5	150	9	31.4
8	3	150	9	37.5
9	0.5	90	5.5	45.2
10	3	90	5.5	50.8
11	1.75	30	5.5	59.5
12	1.75	150	5.5	62.3
13	1.75	90	2	75.4
14	1.75	90	9	58.4
15	1.75	90	5.5	55.4
16	1.75	90	5.5	56.2
17	1.75	90	5.5	54.9
18	1.75	90	5.5	55.8
19	1.75	90	5.5	55.9
20	1.75	90	5.5	56.1

It was found from Table 4 that the regression F-value was 519.3 and the probability P value was 0, representing the significance of the model. The correlation coefficient for this model (R^2) was 0.95.

Table 4
Anova test results for the adsorption processes

Effect	Sum of Squares	DF	Mean Squares	F-value	P-value
Regression	61226.80	10	6122.680	51.9	0
Residual	117.88	10	11.788		
Total	61344.68	20			

Effect of Dose Solution. The effect of *Cassia surattensis* seeds dose on the adsorption of oil content was deliberated in the dose variety of 0.5 g –3 g with oil concentration of 118 mg/g with normal pH value and at room temperature. The consequences in Figure 2 demonstrated that the proportion oil exclusion improved from 22.5% – 58.4% via a cumulative adsorbent amount of 0.5 gm – 3 gm. This was due to the cumulative adsorption places and high relation of oil content to empty place. Snowballing sites of the adsorption and empty site consumed slight results on oil elimination at high adsorbent quantity since of the discovered equilibrium at every squat adsorbate concentration in the wastewater prior to saturation attainment (Agarwal et al., 2016). A similar result was experienced by Ibrahim et al. (2010) throughout the adsorption squalor of lead (II) ions from aqueous solutions via agricultural waste. Therefore, rendering to this trial the best adsorbent dosage values of 2 g – 3 g and 2 g were nominated for *Cassia surattensis* seeds adsorbent to carry out the adsorption experiments.

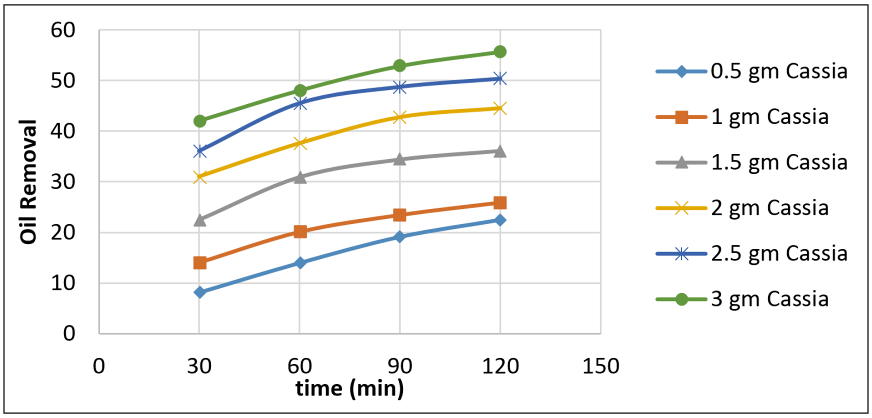


Figure 2. Effect of amount on or elimination in 0.3 mm surattensis seeds size, pH=7.11

Effect of Contact Time and Adsorption Capacity. The consequence of contact time on the adsorption of *Cassia surattensis* seeds is revealed in Figure 3. The figure demonstrates the adsorption in kernels with contact time and reaches to equilibrium after 2.5 h. However, the upsurge was comparatively higher throughout the initial 0.5 h. Fast growth in adsorption throughout the initial stage might presume: capable to stand owing to the obtainability of available active places on the hydrocarbon surfaces. The sluggish upsurge at the later stage was a result of the dispersal of organic into the adsorbent holes since the outside places were totally engaged. It was too contingent from Figure 3 that oil adsorption surges with cumulative contact time. A comparable tendency was specified by Shakoor and Nasar (2016) in their studies of methylene blue dye elimination from artificially polluted water by using Citrus limetta peel waste .The values of equilibrium adsorption capacity (i.e. for the contact time of 2.5 h) must have been planned in Figure 4 in contradiction to equilibrium oil concentration aimed at diverse quantity, as labelled on all points. The figure obviously specified that the equilibrium adsorption capacity increased by cumulative *Cassia surattensis* seeds and contact time (Kyzas et al., 2016).

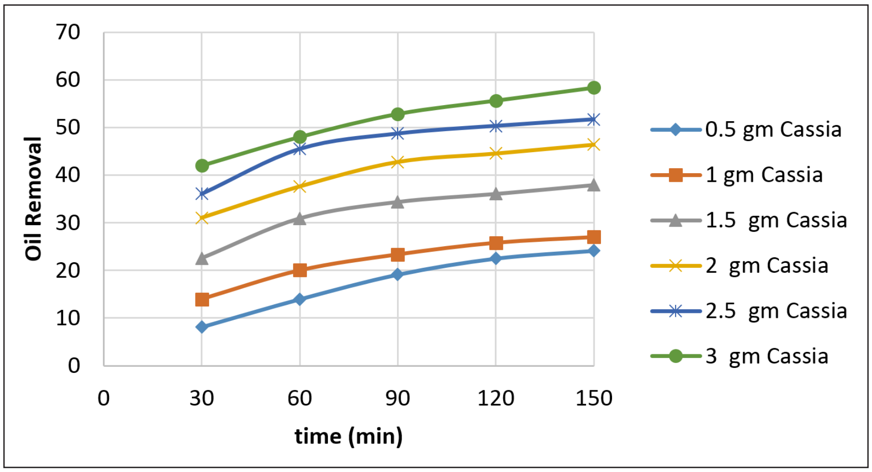


Figure 3. Effect of contact time on oil elimination in 0.3 mm surattensis seeds size, pH=7.11

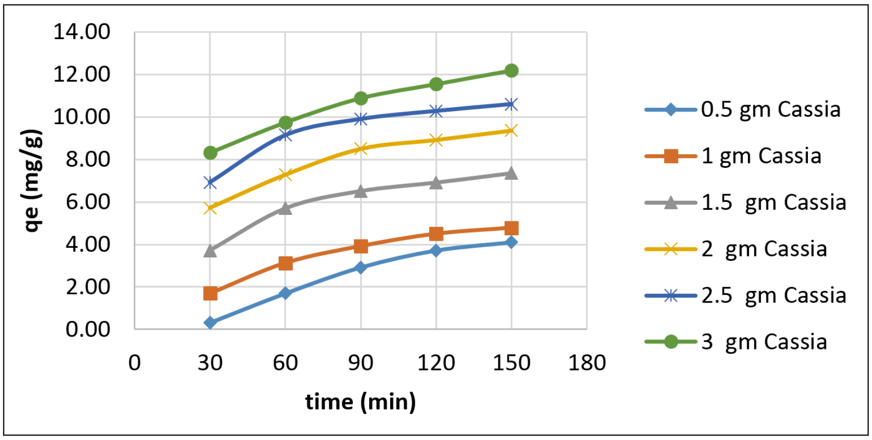


Figure 4. Effect of contact time on adsorption capacity in 0.3 mm surattensis seeds size, pH=7.11

Effect of pH. The result of pH on the adsorption was deliberated by regulating the pH values in the variation of 2.0– 9.0. In these experimentations, the *Cassia surattensis* seeds filling were reserved at 2 g with constant temperature. Figure 5 specifies the effect of pH on the elimination of hydrocarbons compounds on *Cassia surattensis* seeds from PW . The oil exclusion was 30.25 % at pH 9.0, while it was 79.4% for adsorbent at pH 2.0. Intended for pH value of lower than 3.0, the all-out acceptance was reached while is an upheld constant. Supreme adsorption by acidic pH showed that squat pH indicates an upsurge in H+ ions on the adsorbent surface, subsequent to a meaningful robust electrostatic attraction between positively charged *Cassia surattensis* seeds surface and oil content (Omri & Benzina, 2012). The extreme elimination percentage happened at pH 2.0, and hence it was occupied as the best value for more adsorption educations.

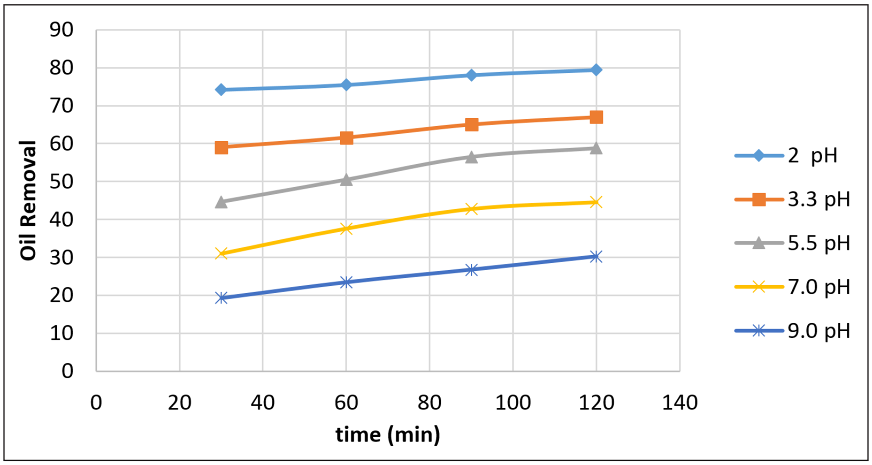


Figure 5. Effect of pH on oil elimination in 0.3 mm surattensis seeds size, 2 gm dose, and 120 min contact time

Adsorption Isotherms

The main adsorption test is a cutting-edge education to clarify the adsorption mechanism. Though, earlier sympathetic adsorption mechanism it is essential to reflect to two opinions: initially the bio-sorbent construction; and then the surface of bio-sorbent possessions. In this assembly, it has the necessity to pierce out that the oil content is a anionic type. On the other hand, *Cassia surattensis* seeds contain hemicellulose, lignin , cellulose and silica as its major components, with other minor components , such as crude protein, fats, and waxes (Majeed et al., 2013). Modification of *Cassia surattensis* seeds with hydrochloric acid removes lignin, silica and additional scums from the seeds surface therefore the skimpy chemical reaction was useful to collect similar – *OH*. In this study, the elimination of oil content through adsorption on *Cassia surattensis* seeds was originated to be fast at the early retro of contact time and to develop slow and become inactive with increase in time. The adsorption was strongly pH-dependent as shown in Figure 6 and Figure 7.

The organic content was sufficiently adsorbed at pH value of between 2.0 and 9.0. The intraparticle dispersal presented a film diffusion to control the sorption kinetics. However, the intraparticle diffusion was not the dominating mechanism (Chowdhury et al., 2011) . The adsorption data for *Cassia surattensis* seeds was suited into the Langmuir and Freundlich isotherm equations. Figure 6 shows that the Langmuir isotherm of the oil content maximum adsorption capacity of for the whole monolayer coverage was 12.18 mg/g. Meanwhile Figure 7 displays the Freundlich isotherm for oil content and the consistent limits (n= – 0.55, qm=4.2). It seems that the Freundlich model finest hysterics the untried consequences ended the experimental variety by good coefficients of correlation (R^2 > 0.95). This tendency is similar to bags stated in the Karnib's study (Karnib et al., 2014).

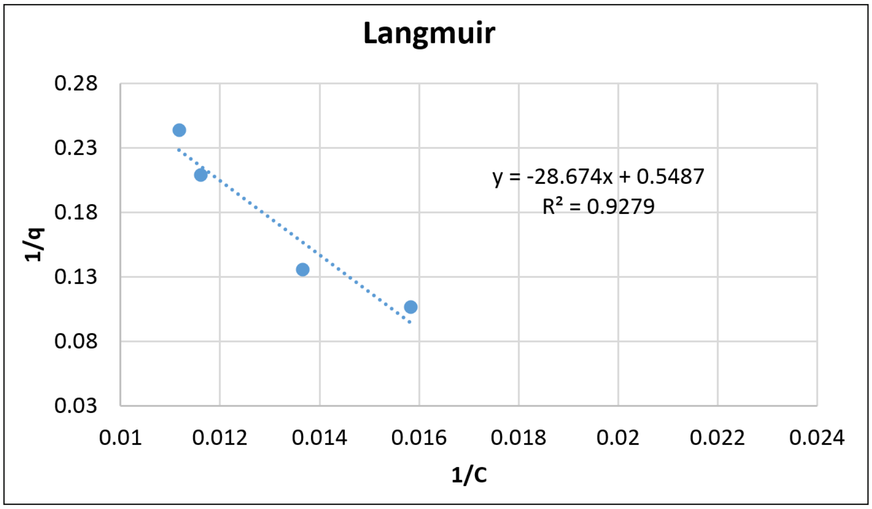


Figure 6. Langmuir isotherm plot for surattensis seeds adsorption of oil content on produced water sample

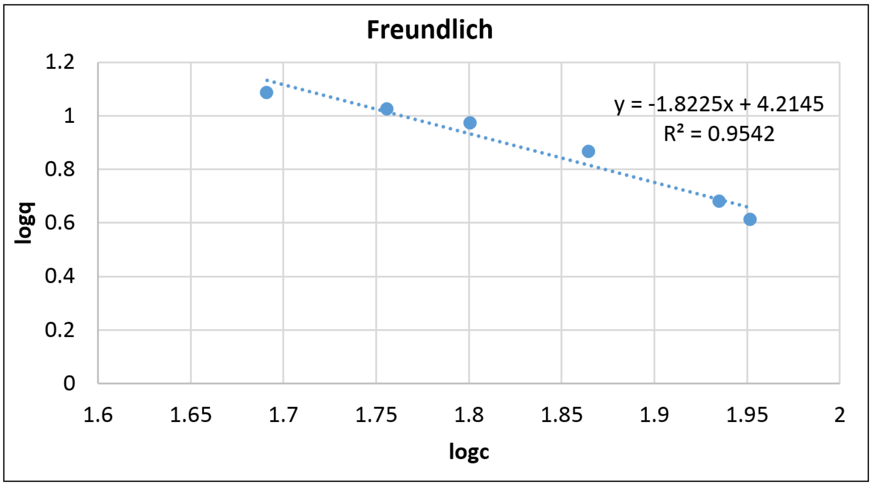


Figure 7. Freundlich isotherm plot for surattensis seeds adsorption of organic content on produced water sample

CONCLUSIONS

The operation of the agricultural substantial as bio-sorbent is an ecological and low cost method. It is a technique to eliminate the oil content in produced water (PW) by accepting the adsorption process of *Cassia surattensis* seeds on oil content and additional wastes. The consequence of many limits on the organic content elimination ,for instance, the adsorbent amount, pH value of the solution and adsorption time have been deliberated. The oil removal was 79.4% at pH 2, 120 min and 2 g amount of *Cassia surattensis* seeds. The comparable Freundlich and Langmuir adsorption isotherm equations were too practical to check the viability of adsorption procedure. However based on correlation constant it was decided that the oil adsorption mechanism in the *Cassia surattensis* seeds system and confirmed with the Freundlich adsorption isotherm model.

ACKNOWLEDGEMENTS

The authors wish to thank University of Al Muthanna and Ministry of Higher Education and Scientific Research, Iraq for funding and supporting this research.

REFERENCES

Agarwal, S., Tyagi, I., Gupta, V. K., Ghasemi, N., Shahivand, M., & Ghasemi, M. (2016). Kinetics, equilibrium studies and thermodynamics of methylene blue adsorption on Ephedra strobilacea saw dust and modified using phosphoric acid and zinc chloride. *Journal of Molecular Liquids, 218*, 208-218.

Aljeboree, A. M., Alshirifi, A. N., & Alkaim, A. F. (2017). Kinetics and equilibrium study for the adsorption of textile dyes on coconut shell activated carbon. *Arabian Journal of Chemistry, 10*, S3381-S3393.

Bagheri, M., Roshandel, R., & Shayegan, J. (2018). Optimal selection of an integrated produced water treatment system in the upstream of oil industry. *Process Safety and Environmental Protection, 117*, 67-81.

Bakke, T., Klungsøyr, J., & Sanni, S. (2013). Environmental impacts of produced water and drilling waste discharges from the Norwegian offshore petroleum industry. *Marine environmental research, 92*, 154-169.

Bediako, J. K., Wei, W., Kim, S., & Yun, Y. S. (2015). Removal of heavy metals from aqueous phases using chemically modified waste Lyocell fiber. *Journal of Hazardous Materials, 299*, 550-561.

Chowdhury, S., Mishra, R., Saha, P., & Kushwaha, P. (2011). Adsorption thermodynamics, kinetics and isosteric heat of adsorption of malachite green onto chemically modified rice husk. *Desalination, 265*(1-3), 159-168.

da Silva, S. S., Chiavone-Filho, O., de Barros Neto, E. L., & Foletto, E. L. (2015). Oil removal from produced water by conjugation of flotation and photo-Fenton processes. *Journal of Environmental Management, 147*, 257-263.

Dahri, M. K., Kooh, M. R. R., & Lim, L. B. (2015). Application of Casuarina equisetifolia needle for the removal of methylene blue and malachite green dyes from aqueous solution. *Alexandria Engineering Journal, 54*(4), 1253-1263.

Dehghani, M. H., Sanaei, D., Ali, I., & Bhatnagar, A. (2016). Removal of chromium (VI) from aqueous solution using treated waste newspaper as a low-cost adsorbent: kinetic modeling and isotherm studies. *Journal of Molecular Liquids, 215*, 671-679.

El Messaoudi, N., El Khomri, M., Bentahar, S., Dbik, A., Lacherai, A., & Bakiz, B. (2016). Evaluation of performance of chemically treated date stones: application for the removal of cationic dyes from aqueous solutions. *Journal of the Taiwan Institute of Chemical Engineers, 67*, 244-253.

Fakhru'l-Razi, A., Pendashteh, A., Abdullah, L. C., Biak, D. R. A., Madaeni, S. S., & Abidin, Z. Z. (2009). Review of technologies for oil and gas produced water treatment. *Journal of Hazardous Materials, 170*(2-3), 530-551.

Fathy, M., El-Sayed, M., Ramzi, M., & Abdelraheem, O. H. (2018). Adsorption separation of condensate oil from produced water using ACTF prepared of oil palm leaves by batch and fixed bed techniques. *Egyptian Journal of Petroleum, 27*(3), 319-326.

Feng, N. C., Guo, X. Y., & Liang, S. (2009a). Adsorption study of copper (II) by chemically modified orange peel. *Journal of Hazardous Materials, 164*(2-3), 1286-1292.

Feng, N. C., Guo, X. Y., & Liang, S. (2009b). Kinetic and thermodynamic studies on biosorption of Cu (II) by chemically modified orange peel. *Transactions of Nonferrous Metals Society of China, 19*(5), 1365-1370.

Foo, K. Y., & Hameed, B. H. (2012). Potential of jackfruit peel as precursor for activated carbon prepared by microwave induced NaOH activation. *Bioresource Technology, 112*, 143-150.

Hassan, A. A., Naeem, H. T., & Hadi, R. T. (2018). Degradation of oily waste water in aqueous phase using solar (ZnO, TiO_2 and Al_2O_3) catalysts. *Pakistan Journal of Biotechnology, 15*(4), 909-916.

Hernández-Francisco, E., Peral, J., & Blanco-Jerez, L. M. (2017). Removal of phenolic compounds from oil refinery wastewater by electrocoagulation and Fenton/photo-Fenton processes. *Journal of Water Process Engineering, 19*, 96-100.

Hosny, R., Fathy, M., Ramzi, M., Moghny, T. A., Desouky, S. E. M., & Shama, S. A. (2016). Treatment of the oily produced water (OPW) using coagulant mixtures. *Egyptian Journal of Petroleum, 25*(3), 391-396.

Ibrahim, M. M., Ngah, W. W., Norliyana, M. S., Daud, W. W., Rafatullah, M., Sulaiman, O., & Hashim, R. (2010). A novel agricultural waste adsorbent for the removal of lead (II) ions from aqueous solutions. *Journal of Hazardous Materials, 182*(1-3), 377-385.

Inyang, M., Gao, B., Yao, Y., Xue, Y., Zimmerman, A. R., Pullammanappallil, P., & Cao, X. (2012). Removal of heavy metals from aqueous solution by biochars derived from anaerobically digested biomass. *Bioresource Technology, 110*, 50-56.

Jafer, A. S., Hassana, A. A., & Naeem, Z. T. (2019). A study on the potential of moringa seeds in adsorption of organic content from water collected from oilfield refinery. *Pakistan Journal of Biotechnology, 16*(1), 23-29.

Jiménez, S., Micó, M. M., Arnaldos, M., Ferrero, E., Malfeito, J. J., Medina, F., & Contreras, S. (2017). Integrated processes for produced water polishing: Enhanced flotation/sedimentation combined with advanced oxidation processes. *Chemosphere, 168*, 309-317.

Karnib, M., Kabbani, A., Holail, H., & Olama, Z. (2014). Heavy metals removal using activated carbon, silica and silica activated carbon composite. *Energy Procedia, 50*, 113-120.

Kumar, S. U., Subramanion, J. L., Sivapragasam, G., Dharmaraj, S., Chen, Y., & Sreenivasan, S. (2014). Standardization and Quality Evaluation of *Cassia surattensis* Seed Extract. *Research Journal of Pharmaceutical, Biological and Chemical Sciences, 5*(5), 355-363.

Kusworo, T. D., Aryanti, N., & Utomo, D. P. (2018). Oilfield produced water treatment to clean water using integrated activated carbon-bentonite adsorbent and double stages membrane process. *Chemical Engineering Journal, 347*, 462-471.

Kyzas, G. Z., Deliyanni, E. A., & Matis, K. A. (2016). Activated carbons produced by pyrolysis of waste potato peels: cobalt ions removal by adsorption. *Colloids and Surfaces A: Physicochemical and Engineering Aspects, 490*, 74-83.

Li, Q., Kang, C., & Zhang, C. (2005). Waste water produced from an oilfield and continuous treatment with an oil-degrading bacterium. *Process Biochemistry, 40*(2), 873-877.

Mousa, K. M., & Al-Hasan, A. A. (2017, October 4). Oilfield produced water treatment by coagulation/flocculation processes. In *The second conference of post graduate researches (CPGR'2017)* (pp. 75-79). Baghdad, Iraq.

Majeed, K., Jawaid, M., Hassan, A., Bakar, A. A., Khalil, H. A., Salema, A. A., & Inuwa, I. (2013). Potential materials for food packaging from nanoclay/natural fibres filled hybrid composites. *Materials and Design, 46*, 391-410.

Maurya, I. C., Singh, S., Gupta, A. K., Srivastava, P., & Bahadur, L. (2018). Dye-Sensitized Solar Cells Employing Extracts from Four Cassia Flowers as Natural Sensitizers: Studies on Dye Ingredient Effect on Photovoltaic Performance. *Journal of Electronic Materials, 47*(1), 225-232.

Naeem, H. T., & Hassan, A. A. (2018). Effectiveness & economy of sawdust wood adsorbents in removing anionic dyes of aqueous solutions. *Pakistan Journal of Biotechnology, 15*(2), 311-20.

Naeem, H. T., Hassan, A. A., & Al-khateeb, R. T. (2018). Wastewater-(Direct Red Dye) Treatment-Using Solar Fenton Process. *Journal of Pharmaceutical Sciences and Research, 10*(9), 2309-2313.

Omri, A., & Benzina, M. (2012). Removal of manganese (II) ions from aqueous solutions by adsorption on activated carbon derived a new precursor: Ziziphus spina-christi seeds. *Alexandria Engineering Journal, 51*(4), 343-350.

Owalude, S. O., & Tella, A. C. (2016). Removal of hexavalent chromium from aqueous solutions by adsorption on modified groundnut hull. *Beni-Suef University Journal of Basic and Applied Sciences, 5*(4), 377-388.

Pereira, F. V., Gurgel, L. V. A., & Gil, L. F. (2010). Removal of Zn2+ from aqueous single metal solutions and electroplating wastewater with wood sawdust and sugarcane bagasse modified with EDTA dianhydride (EDTAD). *Journal of Hazardous Materials, 176*(1-3), 856-863.

Pirbazari, A. E., Saberikhah, E., Badrouh, M., & Emami, M. S. (2014). Alkali treated Foumanat tea waste as an efficient adsorbent for methylene blue adsorption from aqueous solution. *Water Resources and Industry, 6*, 64-80.

Rangabhashiyam, S., & Selvaraju, N. (2015). Efficacy of unmodified and chemically modified Swietenia mahagoni shells for the removal of hexavalent chromium from simulated wastewater. *Journal of Molecular liquids, 209*, 487-497.

Shakoor, S., & Nasar, A. (2016). Removal of methylene blue dye from artificially contaminated water using citrus limetta peel waste as a very low cost adsorbent. *Journal of the Taiwan Institute of Chemical Engineers, 66*, 154-163.

Shokrollahzadeh, S., Golmohammad, F., Naseri, N., Shokouhi, H., & Arman-Mehr, M. (2012). Chemical oxidation for removal of hydrocarbons from gas–field produced water. *Procedia Engineering, 42*, 942-947.

Yang, J., Yu, M., & Chen, W. (2015). Adsorption of hexavalent chromium from aqueous solution by activated carbon prepared from longan seed: Kinetics, equilibrium and thermodynamics. *Journal of Industrial and Engineering Chemistry, 21*, 414-422.

Yu, H., Pang, J., Ai, T., & Liu, L. (2016). Biosorption of Cu2+, Co2+ and Ni2+ from aqueous solution by modified corn silk: Equilibrium, kinetics, and thermodynamic studies. *Journal of the Taiwan Institute of Chemical Engineers, 62*, 21-30.

Pertanika J. Sci. & Technol. 27 (4): 2139 - 2152 (2019)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Modelling and Optimization of Torrefied Pellet Fuel Production

Ahmad Hafizi Awang*, Abdulhalim Abdulrazik, Azuin Mad Noor and Aainaa Izyan Nafsun

Faculty of Chemical and Process Engineering Technology, Universiti Malaysia Pahang, 26300 UMP, Gambang, Pahang, Malaysia

ABSTRACT

Torrefaction is a thermal process to convert biomass into a coal-like material, which has better fuel characteristics than the original biomass. Torrefied biomass has more energy density and hydrophobic which is superior quality for handling and storage. The objective of this research was to develop a simulation model of the torrefied pelletization process from empty fruit bunch (EFB). The process was simulated using ASPEN Plus. Optimization involved a selection of the model option that produced the maximum mass yield and minimum energy requirement, with a converged base case simulation as a starting point. Torrefied biomass pellet offered coal-like properties such as high heating value, brittle, high bulk energy density and more hydrophobic. These properties could potentially avoid costly power plant modifications. On the other hand, Malaysia has issued National Biomass Strategy 2020 with target to solve the problem of under-utilized biomass in this country. Base model was based on previous study. For optimization of mass yield and overall energy consumption, six model options of design configurations were analysed. Design model 0 was used as the base model. For design model 1, flue gas from combustion reactor was channelled to torrefaction reactor. For design model 2, flue gas from combustion reactor was split to dryer and torrefaction reactor. For design model 3, combustion reactor was removed. For design model 4, flue gas was channelled to dryer reactor without combustion reactor. For design model 5, flue gas separator after dryer was removed. Out of five options, results were tabulated for the optimum one. The results showed that the highest mass yield was achieved by simulation Model 5 at 90.76 % and lowest energy requirement was achieved by

ARTICLE INFO

Article history:
Received: 07 January 2019
Accepted: 04 June 2019
Published: 21 October 2019

E-mail addresses:
hafiziawang94@gmail.com (Ahmad Hafizi Awang)
abdhalim@ump.edu.my (Abdul Halim Abdul Razik)
azuinmn@gmail.com (Azuin Mad Noor)
izyan@ump.edu.my (Aainaa Izyan Nafsun)
* Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

simulation Model 4 at 411.336 kW. Optimization result meanwhile had shown that Model 4 was selected because it gave the maximum profitability of RM 72834.45 by considering the yield and the energy consumption simultaneously.

Keywords: ASPEN Plus, empty fruit bunch (EFB), optimization, torrefaction and pelletization (TOP)

INTRODUCTION

Modelling and optimization of torrefied pellet fuel production from biomass have been a subject of interest in these recent years. In the production process, biomass undergoes torrefaction and pelletization stages to make renewable solid fuel. Torrefied biomass is a treated biomass that may be used as replacement of fossil fuel such as coal to generate electricity at coal-fired power plant in Malaysia. Torrefied biomass pellet has offered coal-like properties such as high heating value, brittle, high bulk energy density and more hydrophobic. After the biomass is torrefied it can be pelletized to ease handling and storage of the material. Prins et al. (2006) added that torrefied version of biomass created homogeneity which was an advantage for automated feeds environment.

Majority of scientists believe that human made carbon dioxide (CO_2) and other greenhouse gasses are one of the causes of global warming directly heating the atmosphere. In addition, we have been relying on fossil fuels for energy uses and this situation accelerates CO_2 productions through combustion process. Biomass is known to have the advantage of generating zero net carbon dioxide emission as a carbon neutral material (Novianti et al., 2014). On the other hand, Malaysia has issued National Biomass Strategy 2020 with target solve the problem of under-utilized biomass in this country (Melsson, 2013). Therefore, this study is in line with the national goal to increase the usage of biomass where torrefaction and pelletization of raw biomass are playing key roles. Emphases are given to the aspect of modelling and optimization.

Torrefaction of biomass has been widely commercialized in Canada and European countries (Hein, 2011); there are a few torrefaction and pelletization (TOP) plants have been established for example, the Toppel Energy in Netherland. The plants are well established but still need further improvement. Experiment works cannot be done especially when the plants are operating. Based on the study by Dudgeon (2009), to understand the process, process model simulation like ASPEN Plus is an important tool. Optimization can be conducted easily by simulation at lower cost with less time consumption and labour.

MODEL DEVELOPMENT

The main process in this torrefaction and pelletization is classified into three stages which consist of the crusher for size reduction, drying of the feed and the decomposition (torrefaction) of the feed. The torrefaction and pelletization flow sheet is shown in Figure 1. The Readlich-Kwong-Soave (RKS) cubic equation of state with Boston-Mathias alpha function (RKS-BM) is used as a thermos-physical property method in the torrefaction process. RSTOICH reactor is used for reactor where the kinetics are unknown or unimportant, but stoichiometry and other information are known and RYIELD reactor is used for reactor where the stoichiometry and kinetics are unknown or unimportant, but yield distribution is known. The first stage is the crusher to form a pellet of raw empty fruit bunch (EFB) before transferring to the feed dryer. The cooling process can be done by blowing air through the pellet as they sit in the metal bin (Ciolkosz, 2009). Next is RSTOICH reactor which is used to simulate the drying of EFB feed to reduce the moisture content before torrefaction stage. When the dry EFB feed enters the RYIELD reactor, the decomposition of the dry EFB feed takes place and the feed is then converted into atoms of carbon (C), hydrogen (H), oxygen (O), sulphur (S), nitrogen (N) and ash. Part of crushed EFB and air entered the RSTOICH block for combustion reaction to produce the desired flue gas. Table 1 and Table 2 show the unit blocks used for the torrefaction model and the summary of unit operation. The proximate and ultimate analysis of all the feedstock used in this simulation model is shown in Table 3. 1000 kg/h raw EFB was used in this simulation process.

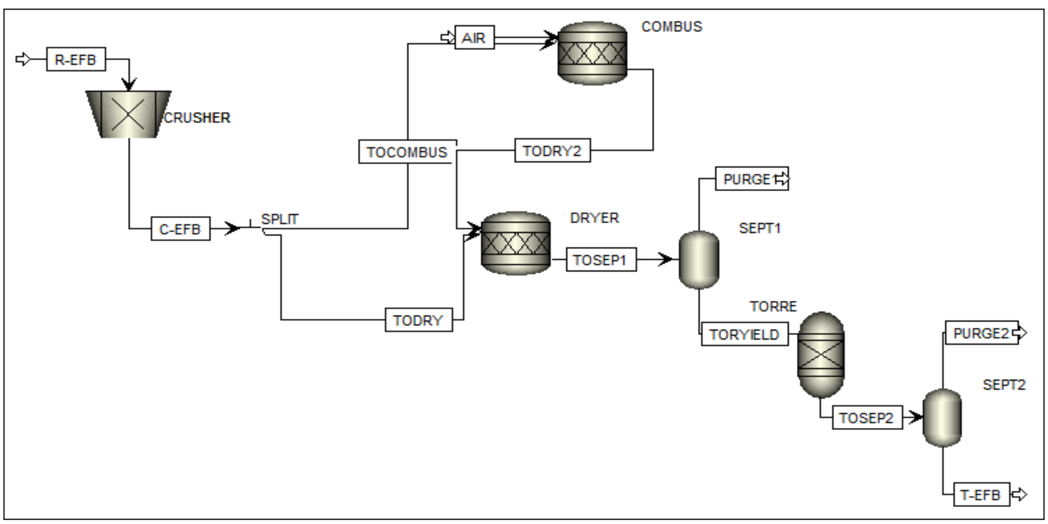


Figure 1. TOP process flow sheet

Table 1
Different unit blocks used for the torrefaction model

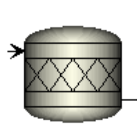
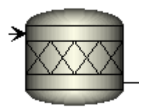
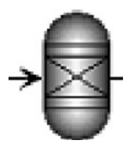
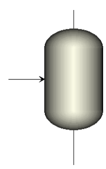
Block Name	Block Type	Block ID	Description	Reaction
COMBUS	RStoic		Combustion reactor - EFB is consumed to produce flue gas that will be used in drying reactor to reduce the energy usage in drying reactor based on mass balance	**Feed → CO_2, H_2O, SO_2, N_2, Energy**

Table 1 *(continue)*

Block Name	Block Type	Block ID	Description	Reaction
DRYER	RStoic		Dryer reactor - convert water in the biomass into moisture based on the extent of reaction.	***Feed*** **→ 0.0555084H_2O**
TORYIELD	Ryield		Yield reactor − convert non-conventional biomass into conventional components based on mass balance	**(Feed) → C, H, O, N, S, ash**
SEPT1 / SEPT2	Flash2		Separation of water and dry biomass	-

Source: Muslim, 2017

Table 2
Summary of Unit Operation

Name	Description	Value Input	Unit
CRUSHER	Crusher		
	- Minimum particle diameter	10	mm
COMBUS	RStoich		
	- Temperature	240	**°C**
	- Pressure	1	**bar**
DRYER	RStoich		
	- Temperature	70	**°C**
	- Pressure	1	**bar**
SEPT1	Flash2		
	- Temperature	160	**°C**
	- Pressure	1	**bar**
TORRE	RYield		
	- Temperature	240	°C
	- Pressure	1	bar

Table 3
Physical and Chemical Properties of EFB used in ASPEN modelling (Olisa, 2014)

	Unit	**Value**
Proximate Analysis		
Moisture	15.00	**% wt**
Volatile	79.82	**% wt**
Fixed Carbon	13.31	**% wt**
Ash	6.87	**% wt**

Table 3 *(continue)*

	Unit	Value
Ultimate Analysis		
Carbon	43.80	**% wt**
Hydrogen	6.20	**% wt**
Oxygen	42.65	**% wt**
Nitrogen	0.44	**% wt**
Sulphur	0.44	**% wt**
Sulphur Analysis		
Pyritic	0.198	**% wt**
Sulfate	0.044	**% wt**
Organic	0.198	% wt

The Chemical Formula of EFB

In this simulation, EFB chemical formula was determined by formula based in Olisa (2014) study that can be written as $C_aH_bO_cN_dS_e$. The EFB chemical formula would be used in the RSTOICH reactor to balance the combustion, drying and torrefaction reaction. The calculation to obtain EFB chemical formula is as follows:

$$C_aH_bO_cN_dS_e$$

$$(Molecular\ weight) \times (Component) = (Ultimate\ analysis) \quad \text{[Eq. 1]}$$

$$(Component) = (Ultimate\ analysis)\ /\ (Molecular\ Weight) \quad \text{[Eq. 2]}$$

C: 12a = 0.438; a = 0.037
H: 1b = 0.062; b = 0.062
O: 16c = 0.427; c = 0.027
N: 14d = 0.0044; d = 0.00031
S: 32e = 0.0044; e = 0.0001375

Therefore, the formula of the EFB can be written as:

$$C_{0.037}H_{0.062}O_{0.027}N_{0.00031}S_{0.0001375}$$

Combustion of EFB in Combustion Reactor

In combustion reactor, EFB is consumed to produce flue gas that will be used in drying reactor to reduce the energy usage in drying reactor. From a study conducted by Olisa (2014), a complete combustion will in theory produce only water vapour, carbon dioxide and other by-products. The combustion equation for EFB with oxygen thus become:

$$C_{0.037}H_{0.062}O_{0.027}N_{0.00031}S_{0.0001375} + xO_2 = pCO_2 + qH_2O + rSO_2 + sN_2 + Energy \quad \text{[Eq. 3]}$$

Balancing the equation given:

$$
\begin{array}{ll}
\text{C: } 1p = 0.037; & p = 0.037 \\
\text{H: } 2q = 0.062; & q = 0.031 \\
\text{O: } 0.027 + 1x = 2(0.037) + 0.031 + 2(0.0001375); & x = 0.0375083 \\
\text{N: } 2s = 0.00031; & s = 0.000155 \\
\text{S: } 1r = 0.0001375; & r = 0.0001375
\end{array}
$$

Hence, the balanced combustion equation thus become:

$$
\begin{aligned}
C_{0.037}H_{0.062}O_{0.027}N_{0.00031}S_{0.0001375} + 0.0391375O_2 = {} & 0.037CO_2 + 0.031H_2O \\
& + 0.0001375SO_2 + 0.000155N_2 \\
& + Energy \qquad \text{[Eq. 4]}
\end{aligned}
$$

Optimization

Process modelling and optimization problems are generally complex tasks, and hence computer software tools are essential for providing fast, reliable and user-friendly interface (Lam et al., 2011). In this paper, optimization to select the best structural configuration of TOP was modelled as mixed integer linear programming (MILP) with binary variables by considering the energy consumption and mass yield. In order to solve the developed optimization model, the General Algebraic Modelling System (GAMS) software version 24.9.2 was used.

Overall Methodology

The overall methodology for this paper is shown by Figure 2.

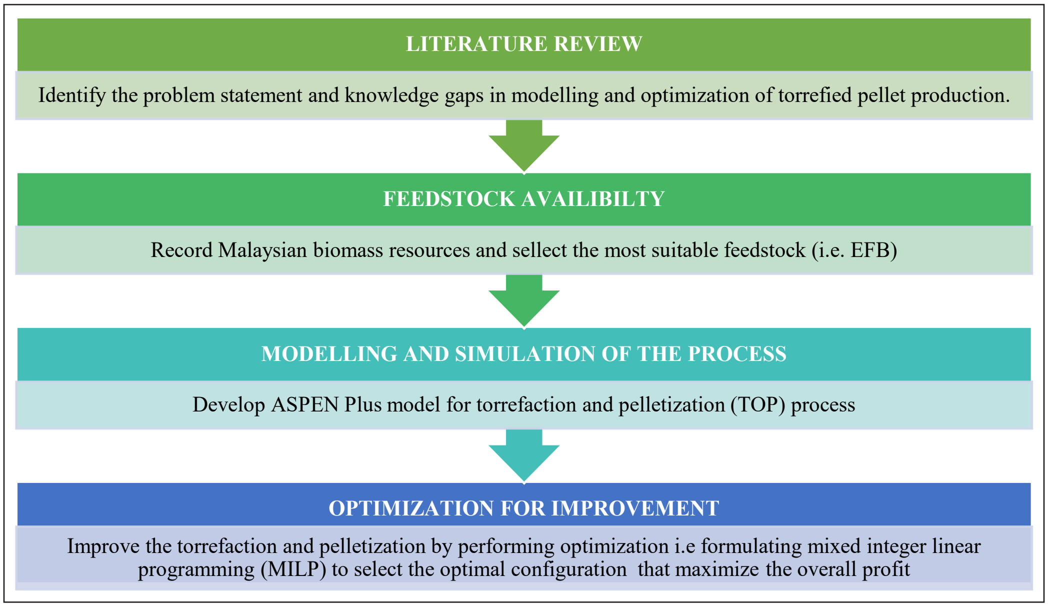


Figure 2. Overall methodology

Diagram for Structural Configurations

For optimization of the overall heat duty, different structure of simulation was studied to obtain the heat duty and compared with other simulation model. Design model 0 was used as the base model. For design model 1, flue gas from combustion reactor was channelled to torrefaction reactor (Figure 3). Therefore, the energy consumption of dryer would increase and for torrefaction would decrease. For design model 2, flue gas from combustion reactor would be split to dryer and torrefaction reactor (Figure 4). Therefore, the energy

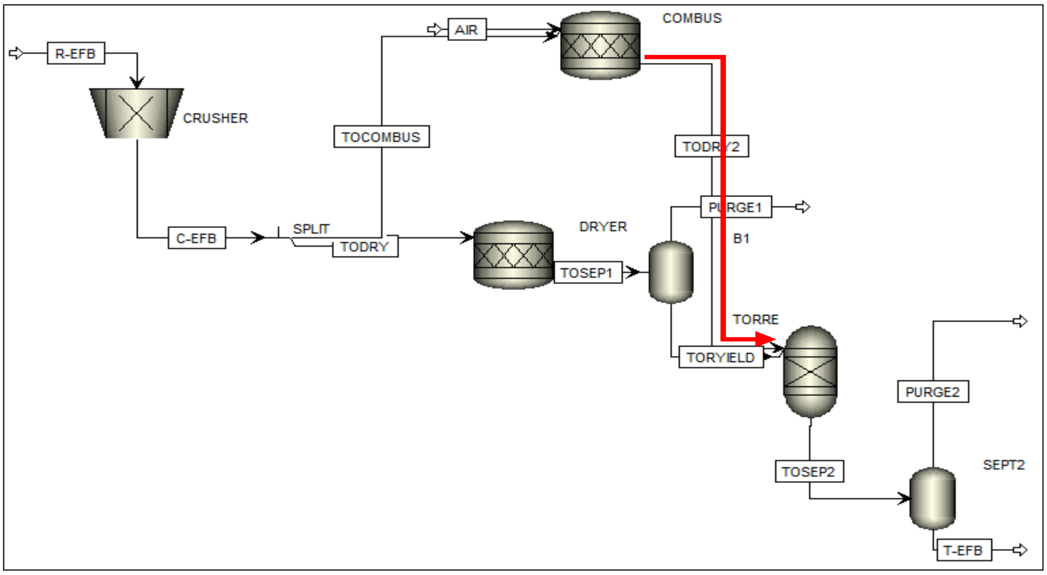


Figure 3. Simulation model 1

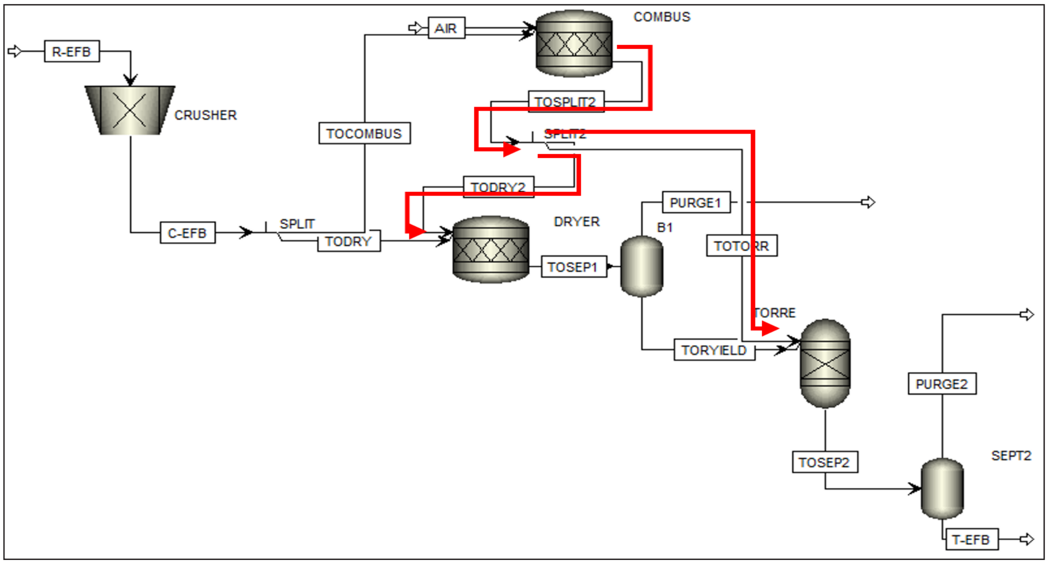


Figure 4. Simulation model 2

consumption for dryer and torrefaction would be reduced. For design model 3, combustion reactor would be removed (Figure 5). Therefore, the energy consumption for dryer would increase with no feedstock EFB used for combustion. For design model 4, flue gas was channeled to dryer reactor without combustion reactor (Figure 6). Therefore, the energy consumption at dryer would be reduced. For design model 5, Separator 1 was removed (Figure 7). Therefore, torrefaction process would include the flue gas from dryer reactor.

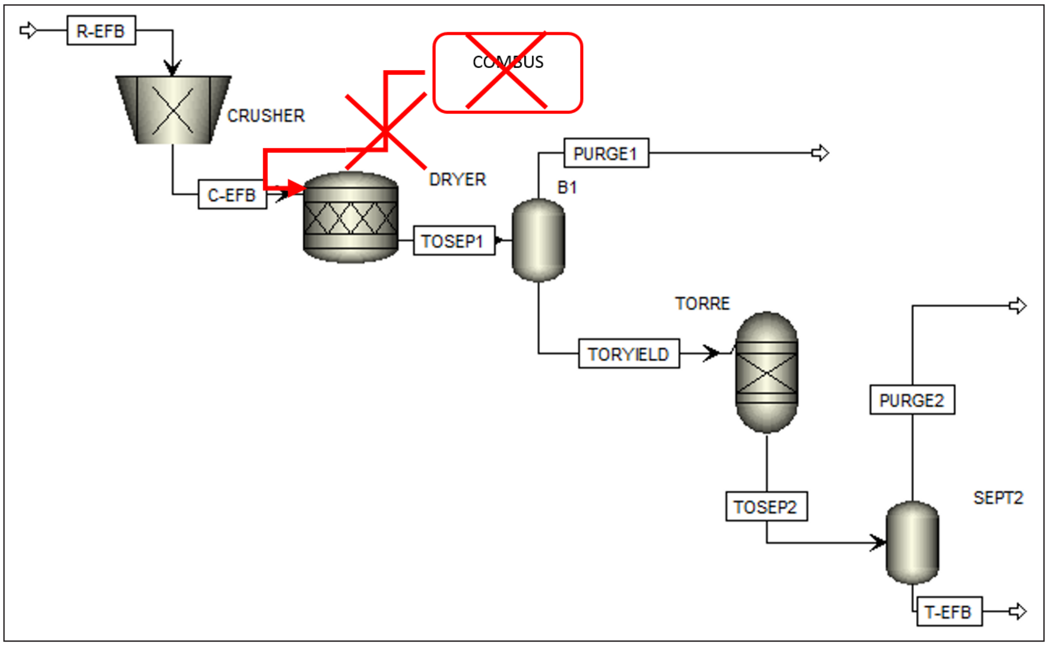


Figure 5. Simulation model 3

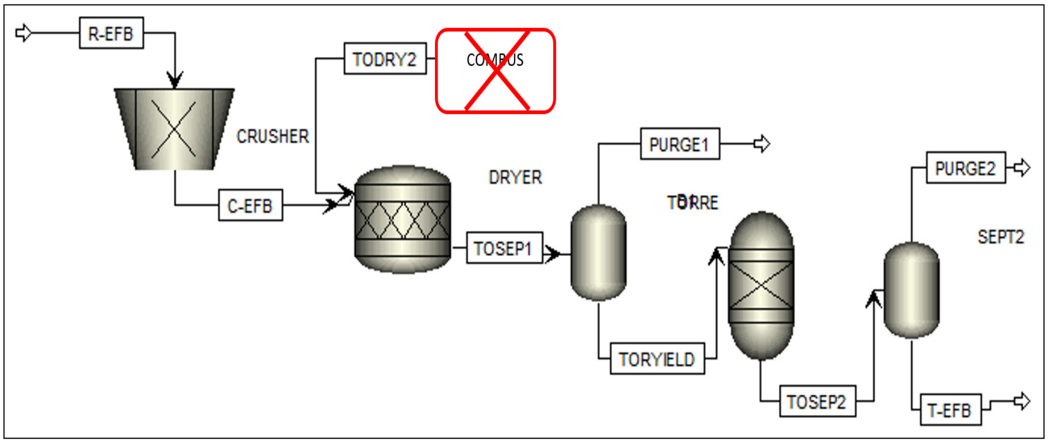


Figure 6. Simulation model 4

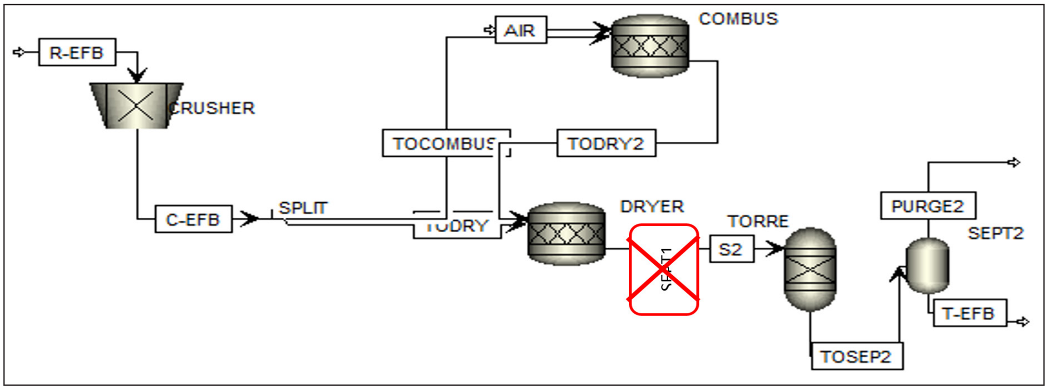


Figure 7. Simulation model 5

Formulation for Optimal Structural Selection

The formulation for the optimal selection was based on the overall profitability in an hour basis. The revenue, energy cost and raw material cost had been considered, while the other associated costs were neglected. For the formulation of the MILP, details as shown below;

Objective function:

Profit = Revenue – Energy Cost – Raw Material Cost

Revenue = Price of Torrefied EFB pellet * Flowrate

Energy Cost = Energy Consumption * Electricity Tariff

Raw Material Cost = EFB Flowrate *EFB Price

x1 = Simulation Model 0
x2 = Simulation Model 1
x3 = Simulation Model 2
x4 = Simulation Model 3
x5 = Simulation Model 4
x6 = Simulation Model 5

$$y = x_1(92010 - 13531.58 - 7000) + x_2(247230 - 53609.98 - 7000) + x_3(170130 - 1134.162 - 7000) + x_4(92010 - 33571.195 - 7000) + x_5(92010 - 12175.55 - 7000) + x_6(272280 - 43025.08 - 7000)$$

[Eq. 5]

Two constraints were identified; i) energy consumption limit, and ii) only one structural configuration would be selected. The energy consumption limit was 510 kW and this was calculated for about 30% of the typical parasitic load (JAREP, 2017). For the selection, 6 binary variables that represent each of the simulation were defined as x_1, x_2, x_3, x_4, x_5, x_6, which the value must be 0 or 1.

c[1]: $457.148x_1 + 1811.148x_2 + 1134.162x_3 + 1295.61x_4 + 411.366x_5 + 1453.55x_6 \leq 510$ [Eq. 6]

c[2]: $x_1 + x_2 + x_3 + x_4 + x_5 + x_6 \leq 1$ [Eq. 7]

All the parameter that used in this formulation were obtained from (JAREP, 2017), and they are as followed:

Parasitic load	= 1.7 MW/h
Energy consumption limit	= 510 kW/h
EFB cost	= RM 7 / ton
TOP EFB cost	= RM 300 / ton
EFB usage	= 600 – 700 ton / day

Electricity Tariff, (Tenaga Nasional, 2018)

TNB Pricing	= RM 29.60 / kWh

RESULT AND DISCUSSION

In biomass energy, torrefaction aims for the production of a fuel that is having improved properties compared to the original biomass. However, this should be achieved without losing too much mass due to release of volatile products during the treatment process. Therefore, the mass fraction is considered the crucial parameter in evaluating a torrefaction process. For the base model, 207.26 kg/h torrefied biomass was produced from 675.8 kg/h EFB feed to torrefaction reactor. The mass fraction yield of the torrefied biomass is defined according to Eq. (5) given as follows:

$$Y_{mass}(\%) = \left(\frac{mass\ after\ torrefaction}{mass\ before\ torrefaction}\right) \times 100 \quad \text{[Eq. 8]}$$

$$Y_{mass}(\%) = \left(\frac{207.26}{675.8}\right) \times 100$$

$$Y_{mass}(\%) = 39.99\%$$

Based on mass yield obtained, high mass reduction occurs due to high composition of volatile composition in feed EFB. High mass loss also occurs during drying stages where high moisture content is reduced before feed into torrefaction reactor.

The overall energy consumption was obtained from ASPEN Plus. Therefore, the overall heat duty for this base model simulation process was 457.1475 kW.

Simulation Result

The overall energy consumption for all simulation studied is compared as shown in Table 4.

Therefore, by comparing the total heat duty for different simulation model, model 4 has the lowest overall heat duty followed by the base model (Figure 8).

From Table 4, simulation model 4 has the lower energy usage at 411.336 kW. The difference of simulation model 4 and base model is that the flue gas is still supplied to the dryer reactor even combustion reactor is removed from the simulation model. This model is only applicable for industry that has already installed combustion reactor in their process with flue gas from existing combustion reactor channelled to TOP process dryer reactor.

Table 4
Summary of energy consumption for different simulation model

Unit Operation	**Energy Consumption (kW)**					
	Model 0 (Base model)	Model 1	Model 2	Model 3	Model 4	**Model 5**
COMBUS	-603.07	-603.07	-603.07	-	-	**-603.07**
DRYER	-253.079	-270.38	-261.73	-337.976	-2054.43	**-253.079**
SEPT1	93.0065	151.228	122.132	189.035	1021.21	
TORR	1386.92	2987.04	2186.98	1641.81	1641.81	**3802.82**
SEPT2	-166.63	-453.67	-310.15	-197.254	-197.254	**-493.121**
Total	457.1475	1811.148	1134.162	1295.615	411.336	**1453.55**

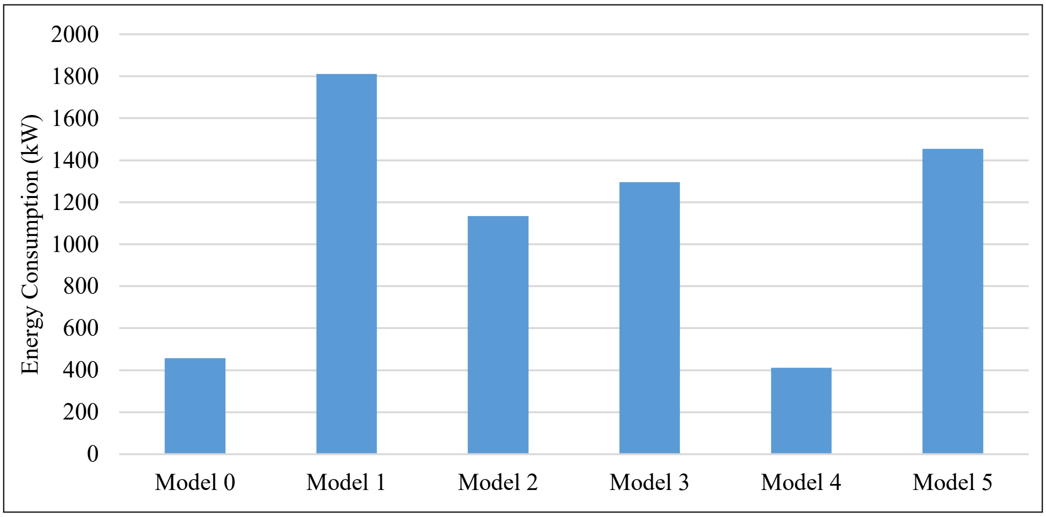


Figure 8. Overall energy consumption at different configuration model

The mass yield is different with different simulation model. The comparison of mass yield of torrefaction reactor is given in Table 5.

Table 5 and Figure 9 show the summary of mass yield for different simulation model and mass yield at different configuration model. For mass yield of torrefied biomass and untreated biomass, simulation model 5 had the highest mass yield at 90.76%. The difference of simulation model 5 and base model is that the flue gas separator after dryer reactor is removed before torrefaction reactor.

Table 5
Summary of mass yield for different simulation model

Simulation	Inlet (kg/h)	Outlet (kg/h)	Y_{mass} (%)
Model 0	675.801	207.255	**30.67**
Model 1	684.7506	564.276	**82.41**
Model 2	680.2753	385.766	**56.71**
Model 3	800	245.345	**30.67**
Model 4	800	245.345	**30.67**
Model 5	675.801	613.345	**90.76**

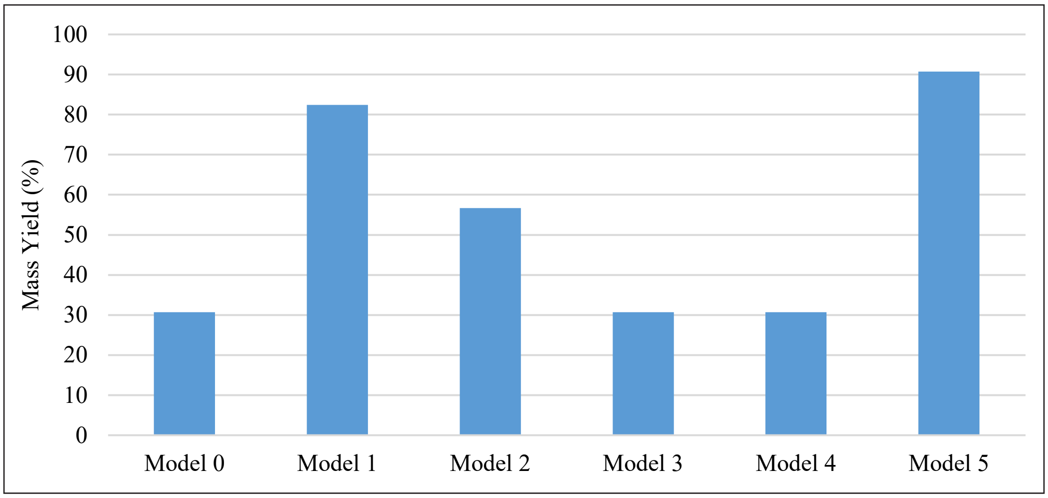


Figure 9. Mass yield at different configuration model

Optimization Result from GAMS

From the GAMS optimization result using CPLEX solver, one simulation model was selected by considering energy consumption and mass yield of TOP simultaneously. Simulation model 4 has been selected to be the best structural configuration with overall profitability of RM 72834.45.

CONCLUSION

In conclusion, the best simulation model should be selected by the simulation that has the lowest overall energy consumption and the highest mass yield of torrefied biomass and untreated biomass. For overall energy consumption, simulation model 4 has the lowest energy usage at 411.336 kW. The difference of simulation model 4 and base model is that the flue gas is still supplied to the dryer reactor even combustion reactor is removed from the simulation model. This model is only applicable for industry that already installed combustion reactor in their process, existing flue gas from the combustion reactor will be channelled to dryer reactor. The simulation model 4 and base model show similar mass yield result of 30.67%. For mass yield of torrefied biomass and untreated biomass, simulation model 5 has the highest mass yield at 90.76%. The difference of simulation model 5 and base model is that the flue gas separator after dryer reactor is removed before torrefaction reactor. Therefore, torrefaction process will include the flue gas from dryer reactor. Based on previous study, it is theoretically impossible for mass yield to be higher than 70% where at least 30% of biomass is a volatile matter that will be removed during torrefaction process (Bergman, 2005). The comparison between simulation model 5 and base model in overall energy consumption is that simulation model 5 is using 1453.55 which is 996.4025 kW higher than the base model. Therefore the best simulation in this paper is simulation model 4 with the lowest overall energy consumption at 411.336 kW and acceptable mass yield at 30.67%. From the optimization result with GAMS and with consideration of both factors, simulation model 4 was selected to be the best structural configuration which can be possibly implemented in any thermal based biomass pre-treatment project.

ACKNOWLEDGEMENTS

This work was supported under the Project Code RDU1703170 and the authors would like to express their gratitude to Universiti Malaysia Pahang for their technical assistance and support in this work.

NOMENCLATURE

%	Percent
%wt	Weight percent
°C	Degree Celsius
bar	Bar (Pressure)
C	Carbon
CO_2	Carbon dioxide
EFB	Empty fruit bunch
GAMS	General Algebraic Modelling System
H	Hydrogen
H_2O	Water
kg/h	kilogram per hour
kW	kilowatt
kW/h	kilowatt per hour
MILP	Mixed integer linear programming
MW/h	Megawatt per hour
N	Nitrogen
O	Oxygen
S	Sulphur
SO2	Sulphur dioxide
TOP	Torrefaction and pelletization
Y_{mass}	Mass yield

REFERENCE

Bergman, P. (2005). *Combined torrefaction and pelletisation (TOP) process* (Report No. ECN-C-05-073). Energy research Centre of the Netherlands ECN, Petten. Retrieved November 22, 2018, from https://publicaties.ecn.nl/PdfFetch.aspx?nr=ECN-C--05-073.

Ciolkosz, D. (2009). *Manufacturing fuel pellets from biomass*. Pennsylvania State University. Retrieved January 15, 2019, from https://extension.psu.edu/manufacturing-fuel-pellets-from-biomass.

Dudgeon, R. (2009). *An Aspen Plus Model of Biomass Torrefaction*. University Turbine Systems Research (UTSR) Fellowship. Retrieved September 24, 2018, from https://deslide.net/viewdoc.html?utm_source=an-aspen-plus-model-of-biomass-tor-ref-action-biomass-gases.

Hein, T. (2011, August 22). *Biomass Torrefaction Technologies*. Retrieved January 15, 2019, from https://www.canadianbiomassmagazine.ca/torrefaction/biomass-torrefaction-technologies-2728

JAREP. (2017). Jengka Advance Renewable Energy Plant. *Personal communication with JAREP operator during industrial attachment*

Lam, H. L., Klemeš, J. J., Kravanja, Z., & Varbanov, P. S. (2011). Software tools overview: Process integration, modelling and optimisation for energy saving and pollution reduction. *Asia-Pacific Journal of Chemical Engineering, 6*(5), 696-712.

Melsson, B. (2013). *National Biomass Strategy 2020: New wealth creation for Malaysia's biomass industry.* Cyberjaya, Malaysia: Agensi Inovasi Malaysia (AIM).

Muslim, B. (2017). Torrefied Biomass Gasification: A Simulation Study by Using Empty Fruit Bunch. In *MATEC Web of Conferences* (Vol. 131, pp. 1-6). Les Ulis, France: EDP Sciences.

Novianti, S. (2014). Upgrading of palm oil empty fruit bunch employing hydrothermal treatment in lab-scale and pilot scale. *Procedia Environmental Sciences, 20*, 46-54.

Olisa, Y. P. (2014). Utilization of palm empty fruit bunch (PEFB) as solid fuel for steam boiler. *European Journal of Engineering and Technology, 2*(2), 1-7.

Prins, M. J., Ptasinski, K. J., & Janssen, F. J. (2006). More efficient biomass gasification via torrefaction. *Energy, 31*(15), 3458-3470.

Tenaga Nasional. (2018, December 14). *Electricity Tariff Adjustment In Peninsular Malaysia for 1 January 2019 to 30 June 2019*. Retrieved March 12, 2019 from https://www.tnb.com.my/announcements/electricity-tariff-adjustment-in-peninsular-malaysia-for-1-january-2019

Pertanika J. Sci. & Technol. 27 (4): 2153 - 2172 (2019)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Techniques and Algorithms for Selection of Number and Locations of Temperature Sensors for Greenhouse

Suman Lata* and Harish Kumar Verma

Department of Electrical and Electronics Engineering, School of Engineering and Technology, Sharda University, India

ABSTRACT

Accurate measurement of spatial distribution of temperature and other micro-climatic parameters inside a greenhouse is important for their monitoring and control. To that end, suitable sensors in adequate number need to be appropriately distributed inside the greenhouse. Two new techniques, namely, Equal Temperature-Step (ETS) and Equal Segment-Area (ESA) techniques are proposed here for the selection of an optimal number of temperature sensors and their locations in a greenhouse with the objective of minimizing the Average Error (AVE), Root Mean Square Error (RMSE) and Maximum Temperature Error (MTE). These techniques were compared with TAE technique, reported earlier. Computational algorithms for the proposed techniques are also presented. Mathematical model of a typical temperature profile along the length of greenhouse has been developed and used for evaluation of the performance. The study shows that the minima of the three errors did not occur simultaneously for any number and locations of the sensors for TAE method. For ETS and ESA methods, the minima of all the three errors occured for the same number and locations of sensors and a smaller number of sensors needed to be used from error consideration. However, reduction in the errors with increase in the number of sensors was steeper for ETS technique as compared to ESA technique, thereby making ETS the best technique. This work can be readily adopted for the measurement of spatial distribution of any other parameter in a greenhouse.

Keywords: Average error, intelligent greenhouse, maximum error, profile, root mean square error, spatial distribution, wireless sensor network

ARTICLE INFO

Article history:
Received: 18 January 2018
Accepted: 4 June 2019
Published: 21 October 2019

E-mail addresses:
suman.lata@sharda.ac.in (Suman Lata)
hk.verma@sharda.ac.in (Harish Kumar Verma)
**Corresponding author*

ISSN: 0128-7680
e-ISSN: 2231-8526

INTRODUCTION

Growth of plants inside a greenhouse depends on several factors, like temperature of the air inside the greenhouse, humidity, soil moisture, intensity of sunlight, carbon dioxide concentration, nutrient level in soil (Ahonen et al., 2008). But temperature is considered as most important among them. Moreover, the environment inside the greenhouse is often non-uniform and dynamic in nature (Pahuja et al., 2012). Hence, to get a good picture of the microclimate of the greenhouse, both spatial and temporal distributions of the selected parameters are important. As for the temperature, its average value inside the greenhouse is of prime importance in case a single crop is grown. In case multiple crops are being grown simultaneously, the profile of the temperature becomes important. To measure the average value and/or the profile of the temperature, adequate number of sensors appropriately located in the greenhouse will be required. Sensors need to be properly distributed along the length as well as the breadth of the greenhouse to obtain the average temperature and/or the temperature profile within acceptable accuracy limits. This calls for optimization of the number of sensors and their locations along both length and breadth of a greenhouse. Moreover, for the development of a cost-effective wireless sensor network (WSN) based measurement system, the deployment of sensor nodes in an Intelligent Greenhouse (IGH) needs to be considered. While an inadequate number of sensors would yield an incorrect measurement of average temperature and temperature profile inside the greenhouse, increasing their number would unnecessarily increase the cost of sensors and sensor nodes. Further, if, in line with the current trend, WSN based temperature measurement system were to be used, the WSN would also become unduly complex and expensive (Pahuja et al., 2013). In order to resolve these issues, it is essential to develop suitable techniques and computational algorithms for selecting an optimal number of sensors and their locations. No research paper focusing on the techniques or algorithms related to selection of number and locations of sensors in a greenhouse could be found in the literature. However, some papers obliquely related to this subject are briefly reviewed in the following paragraphs.

Balendonk et al. (2010) used low cost wireless sensors to investigate the horizontal distribution of temperature and relative humidity and to evaluate the number of sensors needed to accurately estimate the spatial and temporal climate distribution. Authors performed trials in four commercial greenhouses, for which 100 sensors were used. The sensors were placed *at equal distances*. They concluded that 9 sensors per hectare (±33 m spacing) could measure ΔT and ΔRH without missing a cold or wet spot.

Bendigeri and Mallapur (2014) proposed and simulated an energy efficient node placement algorithm (EENPA) for wireless sensor network. The authors worked on a circular node deployment technique instead of random placement of nodes. The simulation was carried out on Qualnet simulator. Simulation results revealed that the proposed algorithm did optimize the energy consumption in the network.

Pandey and Rizvi (2014) analysed certain strategies of node placement in wireless sensor networks. The authors have, in their paper, discussed static and dynamic positioning of nodes and compared various strategies of node positioning in wireless sensor networks. Both, role-based and objective-based placement strategies were considered in detail.

Ryu et al. (2014) conducted experiments in two greenhouses to investigate the spatial, vertical and temporal variability of ambient environment with two different crops. At every layer of measurement the sensors were equi-spaced on the assumption of spatial symmetry of the environmental conditions. The paper did not target on the optimizing the number and distribution of sensors.

Zorzeto et al. (2014) evaluated the homogeneity of distribution of two environmental parameters, temperature and humidity, using wireless sensors in a 1994 m^2 greenhouse with lettuce cultivation. Three sensors were installed in different positions in the greenhouse for 11 days and their hourly averages were collected. Authors concluded that, to assess the homogeneity of temperature and humidity with high accuracy, the number and the locations of the sensors needed to be considered as per application.

Lamprinos et al. (2015) studied the variability of the temperature and humidity inside a greenhouse of 160 m^2 area. The authors had developed a WSN of six nodes out of which three were sensor nodes, one router, one weather station and one coordinating node. In this paper also the distribution of the nodes considered was random.

Somov et al. (2018) developed an IoT system for an operational greenhouse. For the development of the WSN, authors used WaspMote sensor nodes in a mesh topology. The sensor nodes used were having pH, electric conductivity, solution flow, temperature, photo-synthetically active radiation (PAR), humidity and CO_2 measuring sensors .The greenhouse had two zones zone A and zone B. Zone A was used for seed propagation and Zone B for growing plants. Two nodes per tray were deployed with one node in the beginning and the second at the end of each tray.

Lata and Verma (2017) proposed and successfully investigated two approaches namely equal sensor spacing method and trial and error method for the selection of number and locations of temperature sensors in a greenhouse. The trial and error method proved a superior as it offered better accuracy with fewer sensors.

Two new techniques Equal Temperature-Step (ETS) method and Equal Segment-Area (ETA) method are proposed for optimizing the number and locations of temperature sensors from the consideration of various measurement errors and are compared with the Trial-and-Error (TAE) method reported in Lata and Verma (2017).

For each method, a computational algorithm has been developed for determining the optimal number and locations of the sensors. Performance of each technique is investigated in terms of Percentage Average Error (%AVE), Percentage Root Mean Square Error (%RMSE), and Maximum Temperature Error (ΔT_{max}).

Evaluation of the two new techniques and their comparison with the TAE technique are carried out for temperature measurements along the length of the greenhouse only.

The organization of rest of the paper is as follows: A typical temperature profile inside a greenhouse is considered in the next section, while the section thereafter discusses the parameters used for evaluation of the proposed techniques. Subsequent sections are devoted to the description of principles and the algorithms for the proposed techniques, results of evaluation and discussions thereon. Conclusions and scope of future work are presented in the last section.

MATHEMATICAL MODEL OF TEMPERATURE PROFILE INSIDE GREENHOUSE

The greenhouse and the temperature profile inside it have been taken from Lata and Verma (2017) to facilitate comparison of the proposed new techniques with the trial and error method reported in that reference. Thus the greenhouse considered is of 20 m length and W width. It has a door of 2 m height located at the centre of one of the width- wise walls and an air cooler placed at the opposite wall. A layout of the greenhouse is given in the Figure 1.

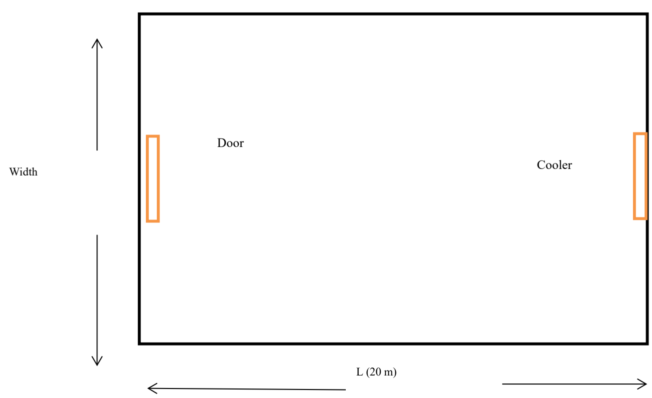


Figure 1. Greenhouse Layout

The temperature profile was developed under following three assumptions:

When the door is closed the temperature just inside the door is 30˚C and temperature outside is 40˚C.

When the door is opened there is a sudden rise in temperature near the door and it decreases exponentially along the length of the greenhouse.

One cooler is kept on opposite side of the door and the temperature at this wall is 20˚C.

The linear component of the profile y_1 is represented by the equation

$$y_1 = mx + c \tag{1}$$

where x = distance from the door along the length of the greenhouse

y_1= temperature varying as a function of x

m= slope of the temperature profile and

c = intercept on the temperature axis.

Applying the assumptions (i) and (iii) listed above, the values of m and c are -0.5˚C/m and 30˚C respectively.

The exponentially component of the temperature profile y_2 is given by

$$y_2 = a\, e^{-x/b} \tag{2}$$

where y_2 = temperature rise at a distance x from the door

a= temperature rise at the door

b= space constant for the exponential decay curve.

As per the assumption (i) outside temperature is 40°C, the temperature near the door on opening would abruptly rise from 30°C to 40°C. Thus, a =10°C. Further it is further it is assumed that space constant ‘b’ for the exponential decay is 2m.

The overall temperature profile inside the greenhouse from the door to the air cooler is obtained by addition of equations 1 and 2, and putting the values of various constants

$$T(x) = y_1 + y_2$$

$$= (mx + c) + (a\, e^{-x/b}) \tag{3}$$

$$T(x) = (-0.5x + 30) + 10e^{-0.5x} \tag{4}$$

This temperature profile is shown in Figure 2.

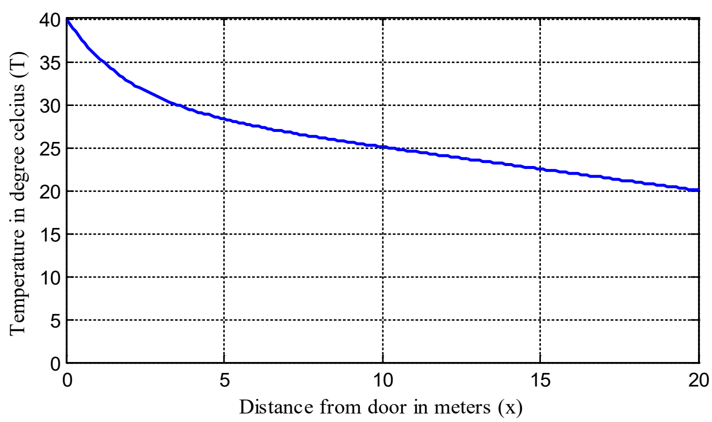


Figure 2. Overall Temperature Profile inside Greenhouse

EVALUATION PARAMETERS

The actual (assumed) temperature profile along the length of GH, T(x), mathematically represented by equation 4, is again shown in Figure 3 (a). For explaining the evaluation process, let the temperature profile curve be broken into four segments (the basis of segmentation is different for the three techniques of selecting the number and locations of sensors). Let a temperature sensor be placed at the middle of each segment and the temperature measured by the sensor be assumed as valid over that segment. The stepped temperature profile so obtained from the sensor measurements, $T_m(x)$, and the variation of square of error with distance from the door are shown in Figure 3 (a) and (b) , respectively. Mathematically, $T_m(x)$, can be expressed by the following relations:

$$T_m(x) = \left.\begin{matrix} T_1 & for\ 0 < x < x_1 \\ T_2 & for\ x_1 < x < x_2 \\ T_3 & for\ x_2 < x < x_3 \\ T_4 & for\ x_3 < x < x_4 \end{matrix}\right\} \quad (5)$$

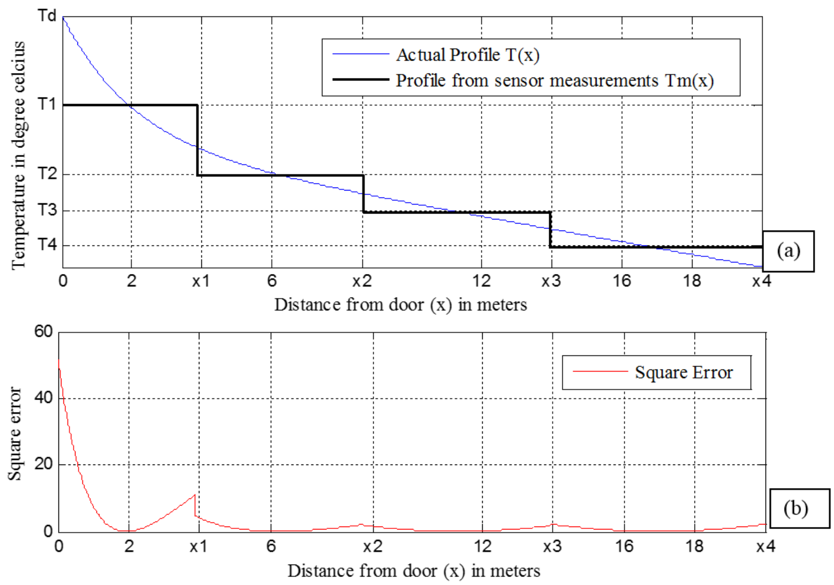


Figure 3. (a) Temperature profiles and (b) Variation of square error with distance from door

For obtaining a temperature profile from sensor data exactly matching with the actual temperature profile, an infinite number of sensors would be required. As per established engineering practices, a small mismatch should always be acceptable. The mismatch can be quantified in terms of certain errors. For that purpose, Percentage average error, Percentage RMS error and Maximum error are used as bases for the evaluation and comparison of the three techniques.

Calculation of Percentage Average Error (%AVE)

The average value of measured temperature along the whole length, T_{avm}, can be obtained by averaging $T_{m(}x)$ over 0 to L. That is,

$$T_{avm} = 1/L \int_0^L T_m(x)\ dx \quad (6)$$

Similarly, the theoretical average value of the temperature along the whole length, T_{avth}, can be obtained by averaging T(x), that is

$$T_{avth} = 1/L \int_0^L T(x)\, dx \tag{7}$$

Therefore, the %AVE in the measured profile can be written as

$$\%AVE = \frac{(T_{avm} - T_{avth})}{T_{avth}} * 100 \tag{8}$$

Calculation of Percentage RMS Error (%RMSE)

The error at a distance x in the measured profile is given by

$$\varepsilon(x) = T_m\ (x) - T(x) \tag{9}$$

The squared error at a distance x will be

$$\varepsilon^2(\mathrm{x}) = [\mathrm{T_m}\ (\mathrm{x}) - \mathrm{T(x)}]^2 \tag{10}$$

So the RMSE along the whole length of GH (L) is given by

$$RMSE = \sqrt{\frac{1}{L}} [\int_0^L [T_m - T(x)]^2\ dx \tag{11}$$

In order to indicate the closeness of the measured profile, $T_m(x)$, to the actual temperature profile, T(x), another unit-less percentage error is defined here. It is taken as the ratio of RMSE to the theoretical value of average temperature along the whole length (T_{avth}), expressed in percentage. Thus

$$\%RMSE = \frac{(RMSE)}{T_{avth}} * 100 \tag{12}$$

Calculation of Maximum Error | ΔT_{max} |

The maximum temperature error | ΔT_{max} | at any point along the profile is yet another indicator of the deviation of the measured profile from the theoretical one, and can be evaluated using the following equation:

$$|\Delta T_{max}| = Max\ [T_m\ (x) - T(x)] \tag{13}$$

DESCRIPTION OF TECHNIQUES

The TAE method proposed in Lata and Verma (2017) and the two newly proposed techniques are described in the following subsections.

Trial-and-Error (TAE) Method

The principle of TAE method proposed in Lata & Verma, 2017 is illustrated in Figure 4 for four sensors (n=4). The minimum number of sensors, n_{min} is assumed as 3 and the minimum spacing between any two sensors is assumed as 2.0m. Therefore, for L=20 m the maximum number of sensors, n_{max} is 20.0/2.0 or 10. The locations of the two extreme sensors are fixed at a distance of $L/2n_{max} =$ 1.0 m from the door and cooler side walls, respectively.

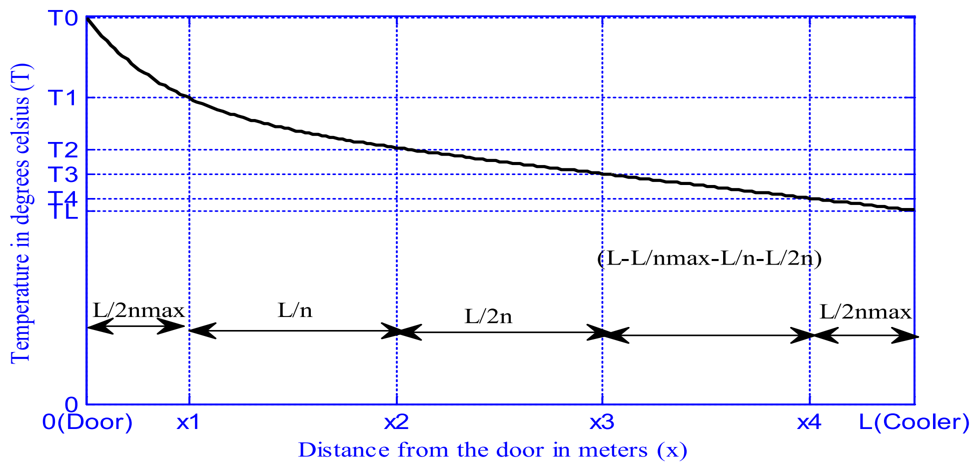


Figure 4. Principle of TAE Method illustrated for Four Sensors

In first trial, the second sensor (S2) from the door side was placed at a distance of L/2n, subject to the condition that it should not be less than 2m from the first sensor S1 and moved further in steps of 1m in the subsequent trials/options. The successive sensors were separated by a distance of L/n. Such trials were made with different number of sensors (3, 4, and so on) and with different placement options. For each option, the average temperature that would be measured by the sensors was calculated and compared with the theoretical average temperature given by equation (7) and the %AVE is determined from equation (8). Similarly, %RMSE and Maximum Error were calculated from equations (11) (12) and (13) respectively.

Equal Temperature-Step (ETS) Method

This method aims at optimizing the number and locations of sensors by segmenting the profile curve on the basis of equal temperature steps. As in the earlier methods, the minimum number of sensors is taken as 3 and the minimum spacing between any two sensors as 2.0m.

ETS Principle

In this approach, whole length of GH (L) is divided into 'n' segments and a sensor is placed at the middle of each segment such that the weighted average of the temperatures measured by the sensors has minimum error with respect to the theoretical average value. The nature of segmentation for n=4 is illustrated in Figure 5, where the first segment is narrowest and the last segment is the widest, and there is a progressive increase in the width in between.

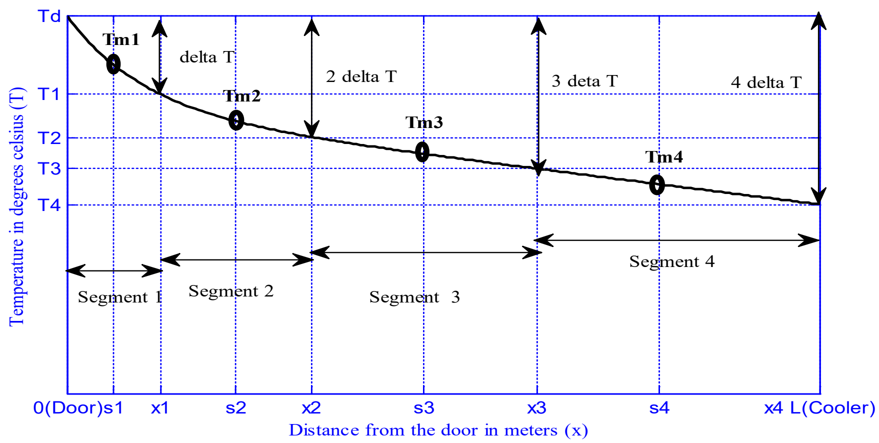


Figure 5. Sensor locations for ETS Method

The principle used here is that the drop from door side temperature (T_d) to cooler side temperature (T_c) is divided into equal steps of ΔT, such that

$$\Delta T = (T_d - T_C)/4 \tag{14}$$

For any general value of n, the relation (14) will become

$$\Delta T = (T_d - T_C)/n \tag{15}$$

The temperatures measured by sensors placed at the middle points of the segments are indicated in the Figure 5 as T_{m1}, T_{m2}, T_{m3} and so on. The weighted average temperature calculated from these values is given by

$$T_{avm} = \frac{1}{L}\sum_{j=1}^{n} T_{mj}\left(x_j - x_{j-1}\right), \qquad x_0 = 0 \tag{16}$$

The percentage error in the measured average temperature can be determined using equation (8), %RMSE using equations (11) and (12) and maximum error using equation (13).These errors are calculated using the above process for different values of 'n' within the limits suggested above, i.e. 3 to 10. The value of 'n' that results in minimum errors is finally selected as the optimal number of sensors to be used. The corresponding locations of the sensors S_1, S_2, S_3 etc. become the optimal locations of these sensors.

Computational Algorithm for ETS

Based on the optimization principle enumerated above, following are the computational steps (algorithm):

Step 1. Calculate the theoretical average temperature from equation (7).

Step 2. Take n=3 (the minimum suggested number of sensors).

Step 3. Calculate temperature step ΔT from equation (15).

Step 4. Calculate the temperatures at the end points of segments from

$$T_j = T_d - j\Delta T, j = \{1,2,3,---- n\} \tag{17}$$

Step 5. For each value of the temperature T_j, j= {1, 2, 3, -----n}, determined in step 4, solve equation 4 numerically to find the corresponding distance x_j (end of j^{th} segment), so that

$$T_j = T(x_j), j = \{1,2,3,--- n\} \tag{18}$$

(18) **Step 6.** Calculate the locations of sensors (distance from door) s_1, s_2 etc. from the equation:

$$s_j = (x_j - x_{j-1})/2 \quad , \qquad j = \{1,2,3,--- n\}, x_0 = 0 \tag{19}$$

Step 7. Calculate the temperatures measured by various sensors from equation 4 by substituting s_j for x, so that

$$T_{mj} = T(s_j), \qquad j = \{1,2,3,--- n\} \tag{20}$$

Step 8. Calculate the weighted average of measured temperatures using equation (16).

Step 9. Calculate %AVE using equation (8).

Step 10. Calculate the profile from the sensor measurements $T_m(x)$ as per equation (5).

Step 11. Calculate RMSE using equation (11) and % RMSE using equation (12).

Step 12. Calculate the maximum temperature error using equation (13).

Step 13. Repeat steps 2 to 12 for n=4, 5 etc. up to the maximum suggested number of sensors or until minima on the modulus of percentage average error versus 'n' curve is achieved.

Step 14. Identify the value of 'n' that gives minimum percentage average error. This is the optimal number of sensors.

Step 15. Take the positions of the sensors corresponding to the optimal value of 'n' from step 6 as the optimal locations of the sensors.

Equal Segment Area (ESA) Method

The method involves optimization of number of locations of sensors by splitting the total area under the profile curve into segments of equal areas.

ESA Principle

The whole area under the temperature versus distance graph along the length of the greenhouse is divided into segments of equal areas and a sensor is placed at the middle of each segment such that the weighted average of the temperatures measured by 'n' sensors has a minimum error with respect to the theoretical average value of the temperature along the length. The limits on the number of sensors are the same as taken in the ETS method. The principle of EAS method on temperature versus distance graph is illustrated in Figure 6 for n=4.

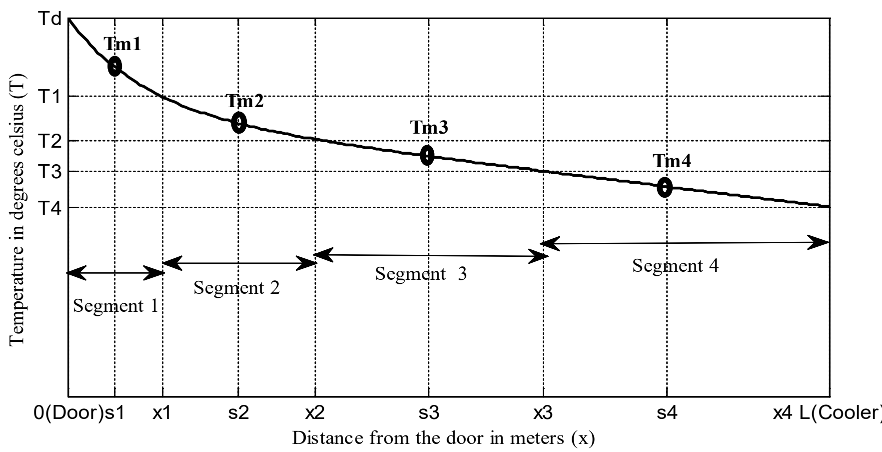


Figure 6. Sensor Locations for ESA Method

The whole area under the graph, A, is split into A_1, A_2, A_3 and A_4 such that

$$A_1 = A_2 = A_3 = A_4 = A/4 \tag{21}$$

For any general value of 'n', relation (21) becomes

Area of each segment = A/n (22)

Total area under the complete curve, A, is calculated as under:

$$A = \int_0^L T(x)\ dx \tag{23}$$

The end points of various segments x_1, x_2, etc. are then successively determined from the following relations:

$$A_1 = \int_0^{x_1} T(x)\ dx = {}^{A}\!/_{4};$$

$$A_2 = \int_{x1}^{x_2} T(x)\ dx = {}^{A}\!/_{4} \qquad \text{and so on} \tag{24}$$

Once the end points of various segments had been determined, the sensors were located at the midpoints along the length of the greenhouse. The temperatures measured by the sensors are indicated in Figure 6 as T_{m1}, T_{m2}, T_{m3} etc. The weighted average temperature was then calculated from these measured values using equation (16), the theoretical average temperature from equation (7) and percentage average error from equation (8). Like the ETS method, the calculations were repeated for different values of 'n' within the suggested limits and optimal number of sensors was identified from the minimum values of the average errors. The corresponding locations of sensors are their optimal locations.

Computational Algorithm for ESA

A computational algorithm for determining the optimal number and locations of sensors based on ESA approach is given below. It is identical to that for ETS approach except at steps 3 and 4.

Step 1. Calculate the theoretical average temperature from equation (7).

Step 2. Take n=3 (the minimum suggested number of sensors).

Step 3. Calculate total area under the curve, A, from equation (23).

Step 4. Calculate end points of various segments x_1, x_2 etc. one by one from the nonlinear equations (24) using an iterative numerical method.

Remaining steps are identical to those of the algorithm for equal temperature-step method.

RESULTS AND DISCUSSIONS

Three techniques TAE, ETS and ESA have been evaluated in respect of the Average, RMS and Maximum errors. The results for each technique along with discussion of the same are given in the following subsections.

TAE Evaluation

The Trial and error process has been applied by varying the number of sensors from 3 to 5. The results for various number of sensors and placement options are given as follows.

Results for Three Sensors (TAE)

In this exercise, three sensors were placed in the greenhouse. The positions of the two extreme sensors (S_1 & S_3) were fixed at 1.0m and 19.0m. In the first option, the second sensor S_2 was placed at a spacing of L/2n=20/6=3.33m from S_1. In subsequent options, this spacing was increased in steps of 1.0m. For example, in option 2, the three sensors are placed at1.0m, 5.33m and 19.0m, respectively, from the door. Various temperature errors are calculated for each option and are given in Table 1.

A graph of these three errors vs. option number is shown in Figure 7. It is observed that %AVE reduces with the increasing option number, but the minima for %RMSE is obtained at option number 5 and the minima for $|\Delta T_{max}|$ occurs at option number 2. However the average error reduces linearly as the option number is increased from 1 to 10.The minima for |%AVE | is for option 10.

Table 1

Placement options along with evaluation errors for three sensors

Option Number	S_1 (m)	S_2 (m)	S_3 (m)	%AVE of TAE	%RMSE of TAE	$\|\Delta T_{max}\|$ of TAE
1	1.0	4.33	19.0	9.0366	15.7056	8.4818
2	1.0	5.33	19.0	7.8167	14.6369	7.5343
3	1.0	6.33	19.0	6.8246	13.8868	8.3082
4	1.0	7.33	19.0	5.9706	13.4205	8.9743
5	1.0	8.33	19.0	5.2005	13.2240	9.5750
6	1.0	9.33	19.0	4.4811	13.2885	10.1316
7	1.0	10.33	19.0	3.7925	13.6011	10.6732
8	1.0	11.33	19.0	3.1227	14.1413	11.1965
9	1.0	12.33	19.0	2.4642	14.8839	11.7093
10	1.0	13.33	19.0	1.8126	15.8000	12.2176

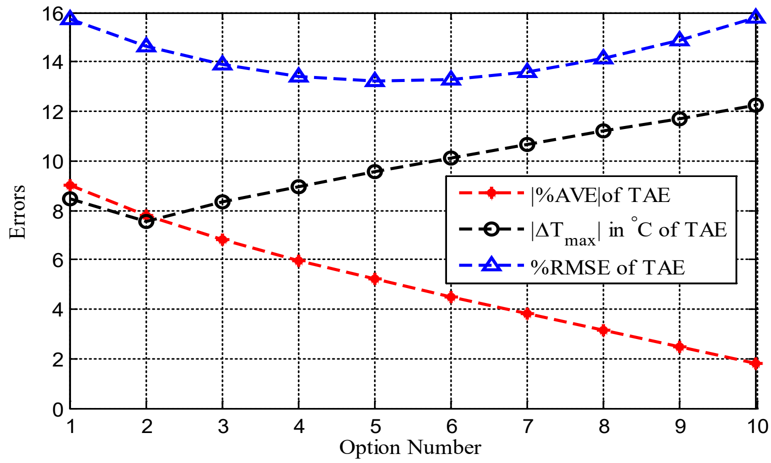


Figure 7. Errors vs. Option number for Three Sensors for TAE

Results for Four Sensors (TAE)

The locations of the two extreme sensors were taken the same as for three-sensor placement options. In the first option, S_2 was placed at a spacing of L/2n=20/8=2.5 m from S_1. This spacing was increased in steps of 1.0 m again. Sensors S_2 and S_3 were separated by a distance of L/n= 20/4 = 5m. Various errors calculated for the 8 options are given in Table 2 and are plotted against the sensor placement options in Figure 8.

Table 2

Placement options along with evaluation errors for four sensors

Option Number	S1 (m)	S2 (m)	S3 (m)	S4 (m)	%AVE of TAE	%RMSE of TAE	$\lvert\Delta T_{max}\rvert$ of TAE
1	1.0	3.5	8.5	19.0	7.6408	10.4051	5.5776
2	1.0	4.5	9.5	19.0	5.9679	9.9195	6.7613
3	1.0	5.5	10.5	19.0	4.5748	9.7979	7.6760
4	1.0	6.5	11.5	19.0	3.3516	9.9633	8.4276
5	1.0	7.5	12.5	19.0	2.2313	10.3646	9.0801
6	1.0	8.5	13.5	19.0	1.1735	10.9632	9.6727
7	1.0	9.5	14.5	19.0	0.1535	11.7255	10.2288
8	1.0	10.5	15.5	19.0	-0.8434	12.6224	10.7628

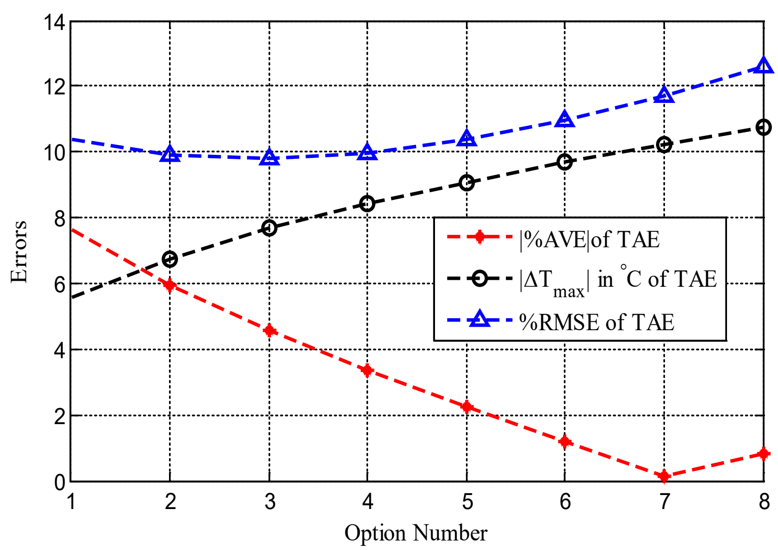


Figure 8. Various Errors vs. Option Number for Four Sensors for TAE

Least %AVE is 0.1535% for option 7, whereas the least %RMSE is 9.7979% for option 3 and the least values of $\lvert\Delta T_{max}\rvert$ is 5.5776 ˚C for option 1. Thus it had been observed that the three errors did not have simultaneous least values for any of the options. For n=4, $\lvert\Delta T_{max}\rvert$ is equal to 5.5776˚C for option 1 and gradually increases to 10.7628˚C for option 8.

Results for Five Sensors (TAE)

The positions of extreme two sensors were again fixed at 1.0 m and 19.0 m. Spacing of S_2 from S_1 is L/2n =20/10=2.0m in the first option, which was increased in the steps of 1.0 m in successive options. S_2, S_3, and S_4 were interspaced by L/n=20/5=4.0m. Various placement options along with the calculated errors given in Table 3 and graphs between various errors and sensor placement options are shown in Figure 9.

Table 3

Placement options along with evaluation errors for five sensors

Option Number	S_1 (m)	S_2 (m)	S_3 (m)	S_4 (m)	S_5 (m)	%AVE of TAE	%RMSE of TAE	$\|\Delta T_{max}\|$ of TAE
1	1.0	3.0	7.0	11.0	19.0	6.2617	7.8754	4.834
2	1.0	4.0	8.0	12.0	19.0	4.3287	7.6993	6.2119
3	1.0	5.0	9.0	13.0	19.0	2.7023	7.9697	7.2445
4	1.0	6.0	10.0	14.0	19.0	1.2619	8.5152	8.0674
5	1.0	7.0	11.0	15.0	19.0	-0.0658	9.2356	8.7633
6	1.0	8.0	12.0	16.0	19.0	-1.3251	10.0760	9.3821

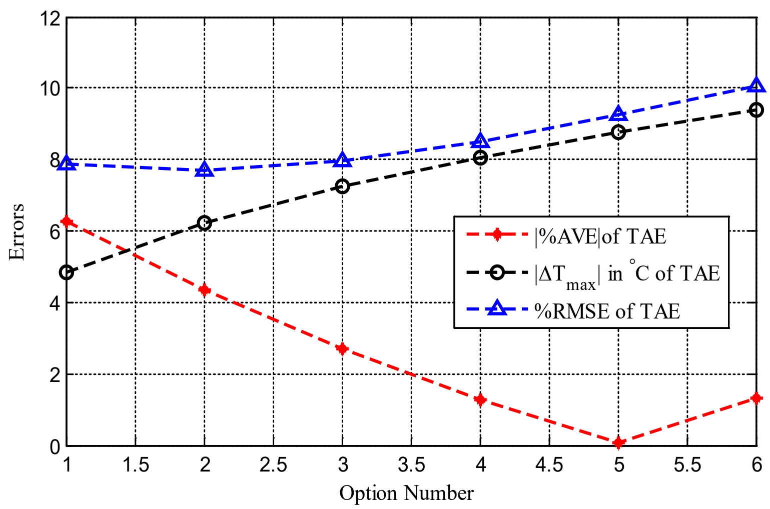


Figure 9. Various Errors vs. Option number for Five Sensors for TAE

It is observed that option 5 gives the least average error of 0.0658%, while option 2 gives the least value of RMS error (7.6993%). Maximum temperature error is lowest for option 1 and increases from option 1 to option 6.

Discussion of Results (TAE)

Based on the above trial-and-error exercise, the least values of %AVE for various number and placement options of sensors are summarized in Table 4 with the corresponding %RMSE and maximum temperature error.

Table 4

Least %AVE for various numbers of sensors along with corresponding %RMSE and $|\Delta T_{max}|$

Number of Sensors	Sensor Locations (Distances from door in m)	Least %AVE of TAE	%RMSE of TAE	$\vert\Delta T_{max}\vert$ of TAE
3	1.0, 13.33, 19.0	1.8126	15.8000	12.2176
4	1.0, 9.5, 14.5, 19.0	0.1535	11.7255	10.2288
5	1.0, 7.0, 11.0, 15.0, 19.0	-0.0658	9.2356	8.7633

It was observed that five sensors, located at 1.0 m, 7.0 m, 11.0 m, 15.0 m and 19 m, yielded a minimum error of 0.0658% in the measurement of average temperature, but the RMSE percentage was 9.2356% for this case, which was not the least. Hence this cannot be considered as the best choice of sensor placement. Least values of %RMSE for various number and placement of sensors are summarized in Table 5 with the corresponding values of %AVE and maximum temperature error. It was observed that five sensors, placed at 1.0m, 4.0m, 8.0m, 12.0m and 19.0m from the door gave %AVE of 4.3287% but corresponding %RMSE was 7.6993% which was quite large. Hence this option is not the best one.

Table 5

Least %RMSE and corresponding %AVE and $|\Delta T_{max}|$ vs. number of sensors

No. of Sensors	Sensor locations (Distances from door in m)	Least %RMSE of TAE	%AVE of TAE	$\vert\Delta T_{max}\vert$ of TAE
3	1.00,8.33,19.00	13.2240	5.2005	9.5750
4	1.00, 5.50, 10.50, 19.00	9.7979	4.5748	7.6760
5	1.00, 4.00,8.00,12.00,19.00	7.6993	4.3287	6.2119

ETS Evaluation

A software program was written in MATLAB as per the algorithm. The results of the computation are summarized in Table 6.

Table 6

Errors vs. number of Sensors for ETS method

No. of Sensors	Locations of Sensors (Distances from door in m)	%AVE of ETS	%RMSE of ETS	$\vert\Delta T_{max}\vert$ of ETS
3	0.8673, 4.4699, 7.2051	-0.5276	7.2831	7.2126
4	0.5829, 2.3284, 6.8088, 15.0637	-0.3375	5.5470	5.2875
5	0.4400, 1.5940, 3.8373, 8.7075, 16.0246	-0.2131	4.4036	4.1604

It may be noted that the percentage error in the average temperature for n = 4 is -0.3375%, root mean square error is 5.5470% and maximum temperature error is 5.2875˚C. For n=5, no significant reduction in these errors has been observed. Thus the optimal number of sensors can be selected as four and the optimal location of these sensors are 0.5829m, 2.3284m, 6.8088m, and 15.0637m from the door. It is important to note that the minima for all the errors occur for the same number and locations of sensors. A graph between various errors and the number of sensors is given in Figure 10.

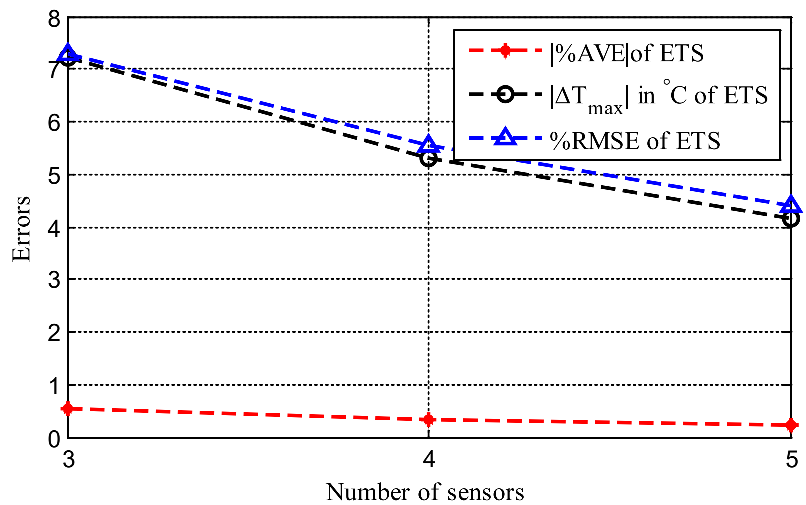


Figure 10. Various Errors vs. Number of Sensors for ETS method

It was observed that the trends for the errors against the number of sensors was similar. For the ETS technique, the minima of all the three errors occured for the same number and locations of sensors. Moreover, as the number of sensors was increased from 3 to 4 in the given situation, there was a significant reduction in each error. But as the number was further increased from 4 to 5, further reductions in the errors were only marginal. Hence instead of five sensors, only four sensors were sufficient. Thus, using ETS technique, cost of one sensor can be saved as compared with the TAE approach.

ESA Evaluation

The proposed ESA algorithm was implemented using MATLAB software. The evaluation was done using same theoretical temperature profile and parameters as used for the ETS and TAE methods. A summary of the results is given Table 7.

Table 7

Errors vs. number of sensors for equal segment-area method

No. of Sensors	Locations of Sensors (Distances from door in m)	%AVE of ESA	%RMSE of ESA	$\lvert\Delta T_{max}\rvert$ of ESA
3	2.6994, 8.7560, 16.0566	-0.9867	7.7687	8.7565
4	1.9440,6.2588,11.2934,16.9787	-0.5798	5.9874	6.6324
5	1.5116,4.8417, 8.6744, 12.8946, 17.5503	-0.3780	4.8483	6.0595

From the Table 7, it can be observed that the least value of all the three errors is corresponding to the same number of sensors, that is, n=5. Again, no significant improvement in %AVE, %RMSE and maximum temperature error is achieved by increasing the number of sensors from 4 to 5. Graphs of all the errors against the number of sensors are shown in Figure 11. For this technique too, the trends for the various errors are similar.

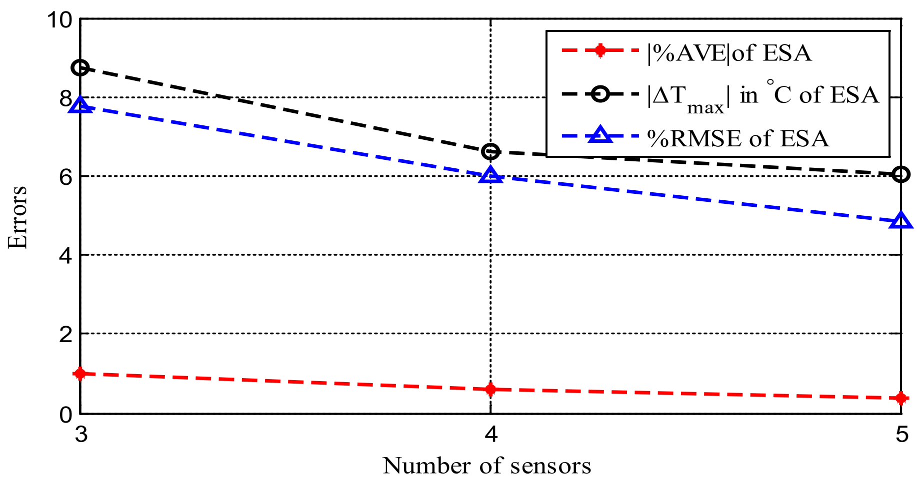


Figure 11. Various Errors vs. Number of Sensors for ESA Method

CONCLUSIONS AND FUTURE WORK

Two new techniques for the selection of an optimal number of temperature sensors and their locations in a greenhouse have been proposed and presented along with computational algorithms. These techniques have been compared with TAE method with respect to various errors. One clear advantage of the newly proposed techniques is that the process of selection of number and location of sensors can be automated using the algorithms developed and reported here. On the contrary, the TAE method which is not based on any logic cannot be automated and needs to be handled manually. Further it is also observed that the minima of the errors do not occur simultaneously for any number and locations of the sensors in TAE Method. On the other hand the minima of all the three errors for the two proposed techniques occur for the same number and locations of the sensors. The reduction in the errors with increase in the number of sensors is not as steep for ESA method as it is for ETS method. Hence it is concluded that the ETS is the best of the three techniques.

The present work has been carried out for the measurement of temperature profile alone and that too along the length of a greenhouse. As a future extension of this work, the proposed techniques can be applied for optimal selection of the number and locations of temperature sensors along the breadth as well. However, the constraint on the minimum number of sensors should to be relooked as per the width of the greenhouse. The resulting arrangement of the sensors will be a two-dimensional matrix. Moreover the techniques can be considered for the measurement of spatial distribution of other micro-climatic parameters in a greenhouse like humidity, luminosity, carbon-dioxide, soil moisture etc., and their performance may be evaluated and compared for different profiles.

ACKNOWLEDGEMENT

The authors wish to gratefully acknowledge the facilities provided by School of Engineering and Technology, Sharda University, Greater Noida, India for carrying the research work reported in this paper.

REFERENCES

Ahonen, T., Virrankoski, R., & Elmusrati, M. (2008, October 12-15). Greenhouse monitoring with wireless sensor network. In *2008 IEEE/ASME International Conference on Mechatronic and Embedded Systems and Applications* (pp. 403-408). Beijing, China.

Balendonck, J., Van Os, E. A., Van der Schoor, R., Van Tuijl, B. A. J., & Keizer, L. C. P. (2010, September 6-8). Monitoring spatial and temporal distribution of temperature and relative humidity in greenhouses based on wireless sensor technology. In *International Conference on Agricultural Engineering-AgEng* (pp.443-453). Clermont-Ferrand, France.

Bendigeri, Y. K., & Mallapur, D. J. (2014). Energy aware node placement algorithm for wireless sensor network. *Advance in Electronic and Electric Engineering*, *4*(6), 541-548.

Lamprinos, I., Charalambides, M., & Chouchoulis, M. (2015, November 15-30). Greenhouse monitoring system based on a wireless sensor network. In *Proceedings of the 2nd International Electronic Conference on Sensors and Applications* (pp. 13-15). Basel, Switzerland.

Lata, S., & Verma, H. K. (2017, November 16-18). Selection of sensor number and locations in intelligent greenhouse. In *2017 3rd International Conference on Condition Assessment Techniques in Electrical Systems (CATCON)* (pp. 58-62). Rupnagar, India.

Pahuja, R., Verma, H. K., & Uddin, M. (2012). Design and implementation of fuzzy temperature control system for WSN applications. *International Journal of Computer Science and Network Security, 12*(11), 1099-1115.

Pahuja, R., Verma, H. K., & Uddin, M. (2013). A Wireless sensor network for greenhouse climate control. *IEEE Pervasive Computing, 12*(2), 49-58.

Pandey, A., & Rizvi, M. A. (2014). Analysis of strategies of placing nodes in wireless sensor networks. *International Journal of Advanced Research in Computer Science and Software Engineering, 4*(2), 398-402.

Ryu, M. J., Ryu, D. K., Chung, S. O., Hur, Y. K., Hur, S. O., Hong, S. J., … & Kim, H. H. (2014). Spatial, vertical, and temporal variability of ambient environments in strawberry and tomato greenhouses in winter. *Journal of Bio-systems Engineering, 39*(1), 47-56.

Somov, A., Shadrin, D., Fastovets, I., Nikitin, A., Matveev, S., Oseledets, I., & Hrinchuk, O. (2018). IoT- enabled greenhouse for plant growth control. *IEEE Pervasive Computing, 17*(4), 65-75.

Zorzeto, T. Q., Leal, P. A. M., Nunes, E. F., & de Araújo, H. F. (2014). Homogeneity of temperature and relative humidity of air in greenhouse. *International Proceedings of Chemical, Biological and Environmental Engineering (IPCBEE), 79*, 25-29.

Pertanika J. Sci. & Technol. 27 (4): 2173 - 2187 (2019)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Three Phase Grid Connected Neutral Point Clamped (NPC) Multilevel Inverter Fed by Two Wind Turbines

Sanusi Alabi Kamilu[1,2]* and Saad Mekhilef[1]

[1]*Power Electronics and Renewable Energy Research Laboratory (PEARL), Department of Electrical Engineering, University of Malaya, 50603 UM, Kuala Lumpur, Malaysia*
[2]*Department of Electrical Engineering, Waziri Umaru Federal Polytechnic, Birnin Kebbi, Kebbi State, Nigeria*

ABSTRACT

This paper presents two gear driven wind turbine generators (WTG) feeding a single three level grid connected NPC inverter. Each component of WTG is made up of wind turbine, 2-mass gear drive, permanent magnet synchronous generator (PMSG), AC-DC-AC power converter. A simple advanced hill climbs search (AHCS) maximum power point tracking algorithm that uses the mechanical power from the Wind turbine was developed to generate proper duty cycle for the control of single stage DC/DC boost converter. The DC link voltages are series interconnected and fed to a sinusoidal pulse width modulation (SPWM) controlled high power inverter. The complete model is simulated using MATLAB/ SIMULINK software under fixed and fluctuating wind speed conditions. Simulation results have shown that WECs exhibit variability in their output power as a result of changes in their prime movers (wind speed).

Keywords: Advanced hill climb search (AHCS), DC-DC boost converter, neutral point clamped inverter, total harmonic distortion (THD), wind energy conversion system, wind turbine generator

ARTICLE INFO

Article history:
Received: 02 December 2018
Accepted: 13 May 2019
Published: 21 October 2019

E-mail addresses:
sanusielect@yahoo.com (Sanusi Alabi Kamilu)
saad@um.edu.my (Saad Mekhilef)
* Corresponding author

INTRODUCTION

Every system in the universe requires energy in order to maintain its existence and the need of this energy has been increasing continuously with the growth in population (Sener AGALAR, 2016). For example, Nigeria is one of the highly ranked producer of crude oil in the world, but its exploration has caused so much devastating effects in the environment leading to serious unrest

ISSN: 0128-7680
e-ISSN: 2231-8526

by the militants. The result of the unrest is the vandalism of so many oil pipe lines leading to serious decline in the electricity generation in the country which amongst other factors cripples the national economy.

The need for alternatives to the fossil fuels owing to global crisis arising from its rapid depletion, green-house gas emission and restrictions to their extensive utilization due to cost, adverse and hazardous environmental effects has led to rapid development in renewable energy (RE) sources which also enhances new technological trends in the field of power electronics converters.

Wind is a typical source of RE and its use has increased drastically over the last decades. The proposed eyesight of “20% Wind Energy by the year 2030” prepared by the U.S Unit of Energy is one of several reasons that embolden wind power development (Veilleux & Lehn, 2010). The advantages of the wind energy are that; it is cost effective, pollution free, and safe (Bull, 2001). It has proven to be the best choice among the RE sources. This paper is focusing on the wind energy conversion systems (WECS) to meet the world energy demand. The wind energy is obtained by the system consisting of a wind turbine, a generator, power electronic converters and control systems. Power electronic converters play a vital role in obtaining high efficiency and good performance of the WECS (Porselvi & Muthu, 2014) and their developments have tremendously enhanced the active and reactive power control at the grid.

However, with this conspicuously growth and the requirement for extra energy demand, single wind turbine generators (WTG) in isolated places are being substituted with arrays of WTGs, forming wind farms. Among the importance of wind farms are; spatial averaging, enhanced capacity and centralized control. Numbers of ways that have been suggested to connect WTGs in a wind farm to the grid can be found in (Chompoo-Inwai et al., 2005; Reidy & Watson, 2005). The interconnection of individual wind turbine generator via a separate DC/AC inverter and with a provision for an optional transformer is one of the simplest method of achieving a wind farm (Reidy & Watson, 2005). Due to the increased number of system units, the failure rate is increased and hence the unreliability. However, interconnection of the DC link voltages has been suggested (Jayasinghe et al., 2010; Nishikata & Tatsuta, 2010; Veilleux & Lehn, 2010). Under this topology, rectified outputs of individual WTGs are connected to a common DC bus either in parallel connection or series connection and fed to a single inverter. Thus, making the system very simple, and the inverter can be placed everywhere without restraint. By that, an optimum site can be selected for wind turbine/generators with a long-distance DC transmission line (Nishikata & Tatsuta, 2010). Another advantage of this topology from economic point of view is that, the cost of system installation and maintenance is reduced and this directly translates to reduction in cost per KWhr of electricity sold to the consumers.

Multilevel inverter represents the driving force in achieving a high voltage, high power in today's RE integration, most especially in wind energy conversion systems. Multi-level inverter (MLI) has the following advantages over two-level inverter; Reduction in Total harmonic distortions (THD), reduced stress on the power switches, modular structure, improved efficiency, smaller output AC filter requirement and lower electromagnetic interference (Do & Nguyen, 2018; Samadaei et al., 2018). In literature, various topologies that have been investigated are; Neutral point clamped (NPC) inverter, diode clamped that was derived from NPC, Flying capacitor, and cascaded H-bridge inverter (Guo et al., 2018; Peng et al., 2010). Cascaded H-bridge inverter has advantages of eliminating capacitor voltage balancing problem and have modularized bridge structure but has serious limited applications in WECS because of the number of DC sources required. Higher number of levels results in reduced harmonics and external filter may not be necessary while adopting flying capacitor topology but is faced with the challenges of complicated switching control and very expensive capacitors. Among the various multi-level inverter topologies, NPC has always been the preferred choice in wind energy applications and has been selected in this paper because of its relatively simple method of control, requiring only one DC source, and high efficiency for fundamental switching frequency.

In order to make optimal use of the available wind energy, variable wind turbine is explored and this can be achieved using doubly fed induction generator (DFIG) or permanent magnet synchronous generator (PMSG). Compared to DFIG, PMSG has the following qualities; its generator has an improved efficiency because of the low rotor losses, generator can be brushless and the grid-fault ride-through capability is less complex (Li & Chen, 2008; Yaramasu et al., 2017), no requirement for excitation and gear drive. Actually, PMSG has the demerit of requiring large converter sizes (100% of rated power compared to 30% of DFIG) but it is still better in terms of overall system efficiency and reliability and is mostly used in modern wind energy conversion systems. Several PMSG based WECS have been proposed (Aliprantis et al., 2000; Nishikata & Tatsuta, 2010; Samanvorakij & Kumkratug, 2013) and different ways of optimizing the aerodynamic power have also been discussed in (Eltamaly et al., 2013).

This paper presents a comprehensive investigation of the proposed topology, which employs series interconnection of the DC link as shown in Figure 1. Uncontrolled rectifier has been used for the AC/DC conversion because of its simplicity. The two-stage voltage source inverter (VSI) with a boost converter is the conventional solution for renewable energy systems (Nguyen et al., 2019). A simple advanced hill climbs search (AHCS) algorithm has been incorporated within each wind turbine to generate the duty cycle necessary for the control of DC/DC Boost power converter and the optimization of aerodynamic power. The AHCS concept is essentially an ''observe and perturb" concept that is used to traverse the natural power curve of the turbine. A single high-power inverter is therefore used to connect the DC link interconnection to the grid via an LC line filter to assist in the reduction of total harmonic distortion (THD).

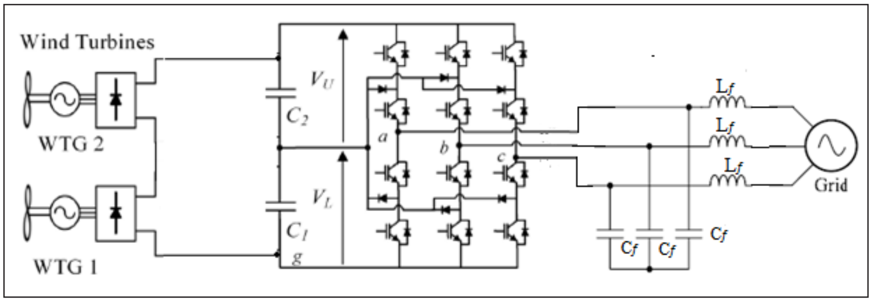


Figure 1. Schematic of the proposed model

MATERIALS AND METHODS

The block diagram of the proposed model is shown in Figure 2. It is made up of two WTGs, each consisting of aerodynamic unit, drive train, AC-DC-AC power converters. The DC link of the two WTGs were series interconnected and fed to an inverter module.

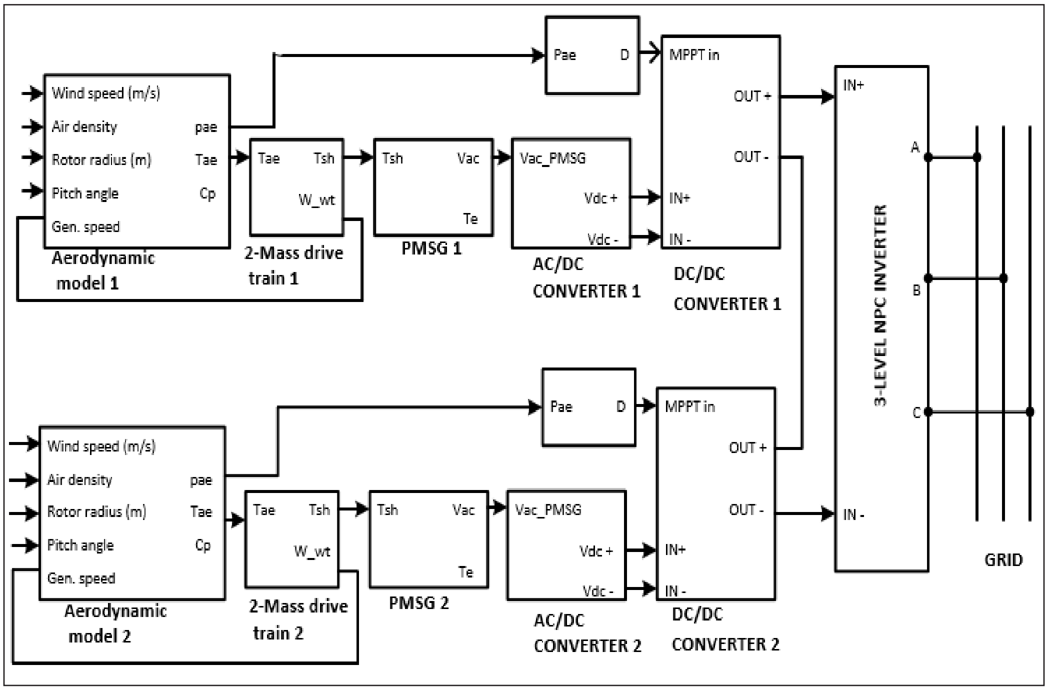


Figure 2. Complete block diagram of the proposed model

Mechanical output power (P_m) of the wind turbine is given by (Samanvorakij & Kumkratug, 2013) as follows;

$$P_m = \frac{1}{2}\rho A v_w^3 C_p(\lambda,\beta) \quad (1)$$

The mechanical torque developed by the wind turbine T_m, and the tip speed ratio (λ) of the wind turbine are given by the following equations;

$$T_m = \frac{P_m}{\Omega} \quad (2)$$

$$\lambda = \frac{\Omega R}{v_w} \quad (3)$$

Where R is the turbine radius and Ω is the angular velocity (rad/sec) of the turbine. There are numbers of fitted equations for power coefficient, C_p, but a more generic one is given as (Eltamaly et al., 2013);

$$C_p(\lambda,\beta) = 0.5176\left\{\left(\frac{116}{\lambda_i}\right) - 0.4\beta - 5\right\}e^{-\frac{21}{\lambda_i}} + 0.0068\lambda \quad (4)$$

Where; $\lambda_i = \frac{1}{\lambda + 0.08\beta} - \frac{0.035}{1+\beta^3}$

In this study, pitch angle, β was set to zero.

Gear train is important in WECs due to the large speed difference between the wind turbine and the generator. In this paper, two mass models were considered. Figure 3 illustrates the schematic of two mass drive train models. The self-damping (D_t) is used to cancel any oscillations that exists in the turbine blade. The generator self-damping (D_g) is the mechanical friction and windage. The mutual damping (D_m) represents balancing dynamics that exists due to varying speeds between the generator rotor and the turbine shaft. Turbine and generator self-damping can be neglected and the resulting mathematical models are given by Aliprantis et al., (2000); and David, (2010).

$$\frac{dw_t}{dt} = \frac{1}{2H_t}(T_m - T_e) \quad (5)$$

$$\frac{1}{w_{ebs}}\frac{d(\theta_t - \theta_r)}{dt} = w_t - w_r \quad (6)$$

$$T_e = K_s(\theta_t - \theta_r) + D_t\frac{d(\theta_t - \theta_r)}{dt} \quad (7)$$

Where;

K_s = Stiffness constant; w_t and w_r are turbine and generator rotor speeds in per units; θ_t and θ_r are turbine and generator angular displacements in rads; T_e = shaft torque.

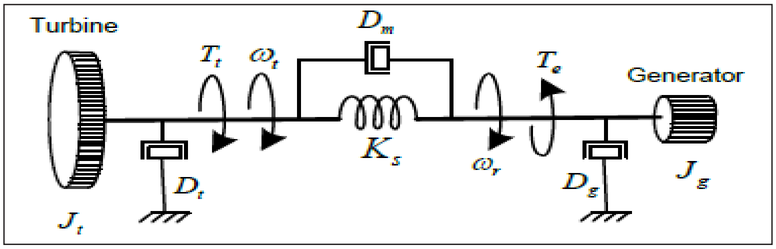


Figure 3. Two-mass drive train model of a wind turbine

Implementation of AHCS Algorithm

Advanced hill climb search MPPT algorithm was used to optimize the aerodynamic power without necessarily requiring wind speed data, it is a sensor-less MPPT controller. The AHCS control algorithm continuously searches for the maximum power of the wind turbine. The optimum power can be derived from WTG without necessarily requiring information about the wind and generator speeds (Eltamaly et al., 2013).

This paper uses a speed control loop concept where the present power is directly compared to the previous power. The MPPT block uses the mechanical power from the Wind turbine model to generate proper duty cycle, which is fed to the DC/DC converter circuit. The time period for which this signal should be ON is generated by the embedded block and is given to the comparator. The comparator output signal is the duty ratio (d) for the boost converter. This block contains an MPPT algorithm which is programed based on the flow chart shown in Figure 4.

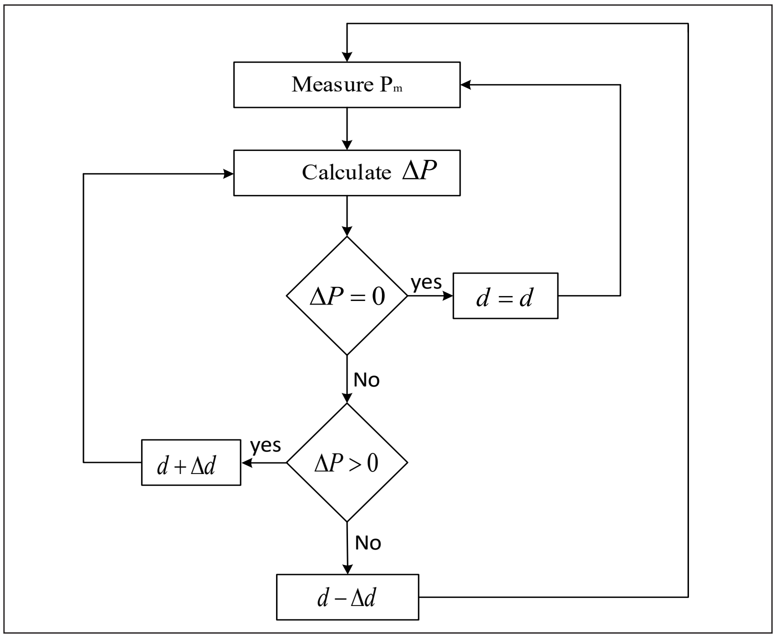


Figure 4. Flow chart of the MPPT algorithm

The three conditions programmed are;

Condition 1: if $\Delta P_m > 200$ then $d = d + \Delta d$

Condition 2: if $\Delta P_m < 200$ then $d = d - \Delta d$

Condition 3: if $\Delta P_m = 200$ then $d = d$

Where;

200 = Reference speed; ΔP_m is the difference between the current power and the previous power; d is the ON time of the duty cycle; Δd is the change in the ON time of the duty cycle. The new power is passed after a delay of sometime and the duty cycle changes till it reaches the optimum condition ($\frac{dP}{dw} = 0$) i.e it extracts the optimum power for the corresponding wind speed.

Design of DC/DC Boost Converter for Continuous Current Mode (CCM)

DC/DC Boost converter is a necessary component for achieving an AC-grid wind turbine. They are used to raise and achieve a stable output DC voltage from a varying input rectifier voltage resulting from wind variations. They also play a vital role in maximizing power generation even at low wind speeds. The objectives here are to design a boost converter with reduced cost, and small settling time. The parameters in Table 1 below are defined by the design specification operating in the Continuous Conduction Mode (CCM).

Table 1
Boost converter design specifications

Design specifications for CCM		
Parameters	**Symbols**	**Values**
Input voltage (Volts)	V_i	100 - 600
Output Voltage in (Volt)	V_o	700
Switching Frequency in (kHz)	f_s	100
Output Power in (KW)	P_o	12.5
The voltage ripple in Volt, $V_r / V_o = 2$ %	V_r	14

The lowest or critical value of inductance for CCM operation is given as;

$$L_{crit} = \frac{R_{max} D_{max}}{2 f_s} (1 - D_{max})^2 \quad for\ D_{min} < \frac{1}{3} \qquad (8)$$

Minimum capacitance value for CCM operation can be obtained using equation (9).

$$C_{min} \geq \frac{D_{max} V_o}{f_s R_{min} V_r} \qquad (9)$$

Detailed information on the design of DC/DC boost converter can be obtained in (Mohamed et al., 2013). Figure 5 shows the complete design schematic of AC-DC, DC/DC converters with gate signal control.

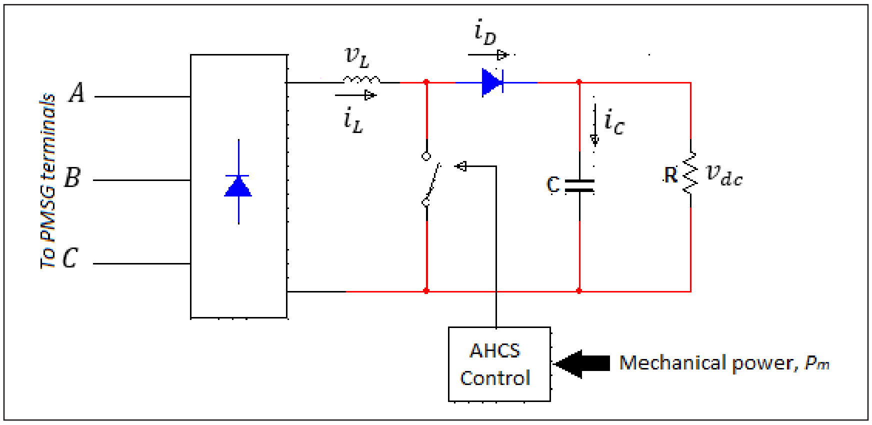


Figure 5. Diode rectifier and DC boost circuit

Inverter performance is influenced by switching strategy, and relate to output voltage harmonic content. There are numbers of inverter switching control techniques, but phase deposition sinusoidal pulse width modulation (PDSPWM) technique has been used in this paper. A high frequency triangular career signal was compared with a low frequency sinusoidal modulating signal to generate gate signals for the control of power switches.

Design of LC filter

In order to reduce grid harmonics injected by the grid connected power converters, LC line filter has been adopted. In most cases, only the output voltage harmonic is explicitly stated. However, the unique value of LC filter cannot be solely specified based on output voltage harmonics. Equations (7)-(9) are derived based on additional criterion that is also based on minimization of reactive power which indirectly minimizes sizes, losses and cost of the filter (Bhende et al., 2011; Dahono & Purwadi, 1995).

$$K = \sqrt{\frac{\left(k^2 - \frac{15}{4}k^4 + \frac{64}{5\pi}k^5 - \frac{5}{4}k^6\right)}{1440}} \tag{10}$$

$$L_f = \frac{V_o}{I_o f_s}\left\{K\frac{V_{dc}}{V_{o,av}}\left[1 + 4\pi^2\left(\frac{f_r}{f_s}\right)^2 K\frac{V_{dc}}{V_{o,av}}\right]\right\}^{1/2} \tag{11}$$

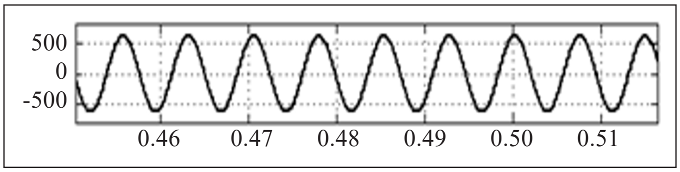
$$C_f = K\frac{V_{dc}}{L_f f_s^2 V_{o,av}} \tag{12}$$

Where; k =Modulation index $\qquad$ f_r = Fundamental frequency =50Hz
V_o = Load voltage (rms) in volts $\qquad$ f_s = Switching frequency = 100 KHz
I_o = Nominal current (A) $\qquad$ L_f = Filter Inductance
$V_{o,av}$ = Total harmonic load voltage= 5% of V_o $\qquad$ C_f = Filter Capacitance

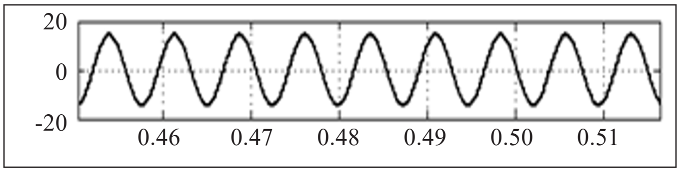
RESULTS AND DISCUSSIONS

The simulations had been carried out using MATLAB/ SIMULINK software at a wind speed of 9m/s and the parameters for the modelling of the drive train, PMSG and wind turbine are as given in Appendix A. Figure 6 illustrates the result for PMSG line voltage and line current without AC-DC-AC converters connected to its terminals. Figure 7 illustrates three phase voltage of PMSG with phase shift of 120 degrees between any two phases.

The simulation results for the control of DC/DC converter using AHCS algorithm is shown in Figure 8. Figure 8(c) illustrates the duty cycle characteristics, by changing d, the



(a)



(b)

Figure 6. PMSG output without AC-DC-AC converters:(a) line voltage; and (b) line current

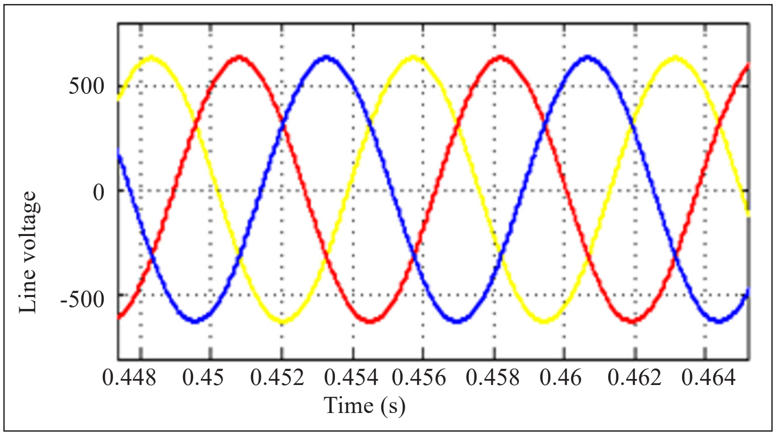


Figure 7. Three phase line-line output voltage of PMSG without power converters

aerodynamic power is changed accordingly and can be optimized. The input and output voltages of the proposed boost converter for continuous current mode at steady state is shown in Figure 9. At a wind speed of 9 m/s, the boost converter input voltage was stepped up from 556.7 Volt to an output voltage of 636.5 Volts. Before the steady state, the DC voltages follow through a transient process which is basically as a result of voltage build up (excitation) process in PMSG.

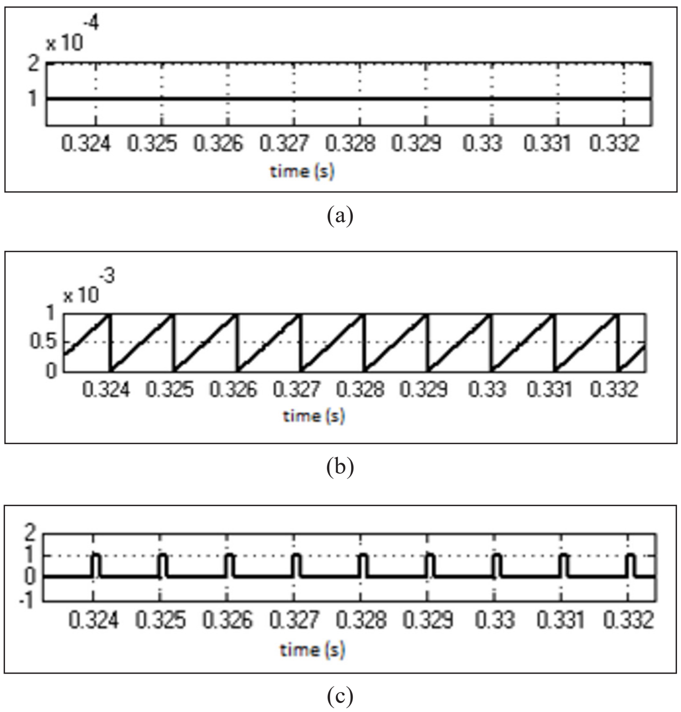


Figure 8. (a) ON time; (b) sawtooth waveform; and (c) switching pulses

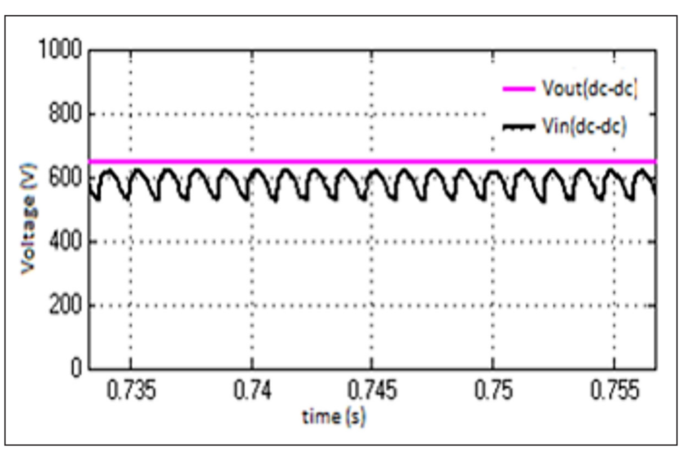


Figure 9. Input and output voltages of DC/DC boost converter

Figure 10 shows the output line voltage of three level NPC inverter. The positive voltage levels corresponds to $0, +\frac{1}{2}V_{dc}, +V_{dc}$, while the negative output voltage level corresponds to $0, -\frac{1}{2}V_{dc}, -V_{dc}$, all totaling to a five levels.

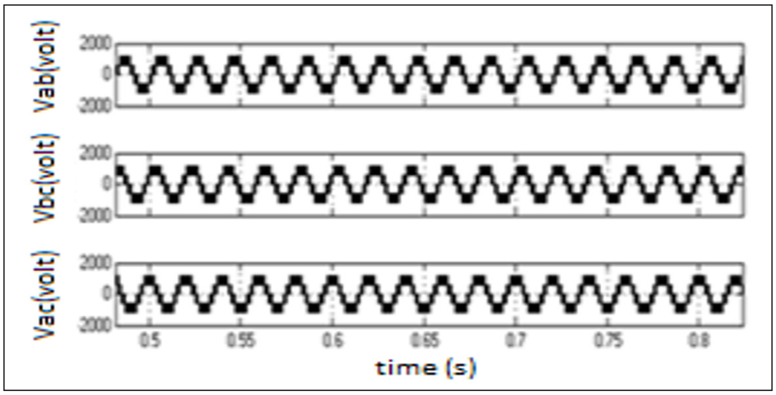


Figure 10. Three phase grid voltage (V_{ab}, V_{ac} and V_{bc}) before filtering

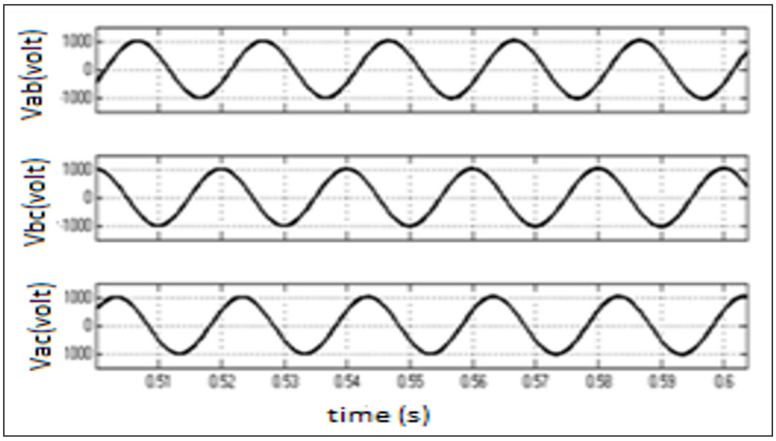
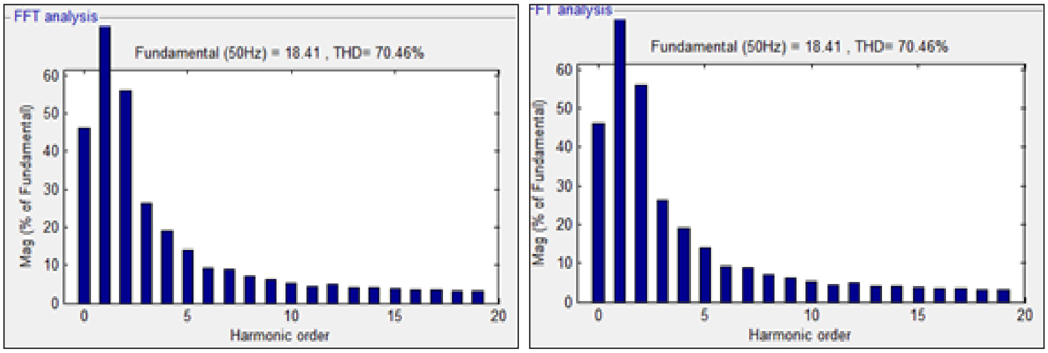


Figure 11. Three phase grid voltage (V_{ab}, V_{ac} and V_{bc}) after LC filtering

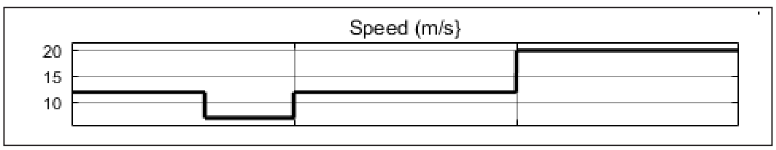
The response is sinusoidal-like and will therefore have a less total harmonic distortion (THD) compared to 2-level inverter. The more the number of levels, the more the response is sinusoidal-like and the less will be the THD. The ouput line to line voltages of the inverter after LC filtering is shown in Figure 11, which is typical of grid voltage. Figure 12 illustrate the voltage THD at the grid before and after the connection of LC line filter which are respectively obtained to be 70.46% and 37.40%.

The results show that the WECs parameters varies in accordance with the wind speed (Figure 13). The effect of this is the variation in operating temperatures which will consequently lead to variation in power losses of the power converters (PC) thereby heightening the unreliability.

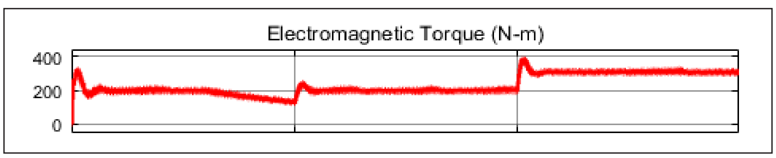


(a) (b)

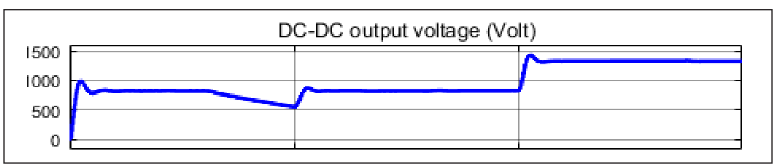
Figure 12. Voltage THD at the grid: (a) before filtering; and (b) after filtering



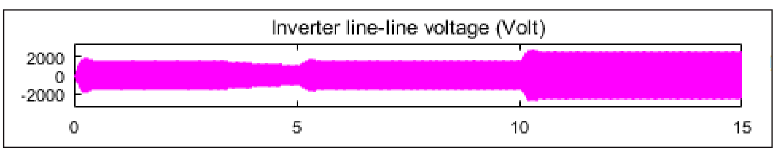
(a)



(b)



(c)



(d)

Figure 13. Impacts of wind speed fluctuations on WECS parameters: (a) wind speed variations; (b) electromagnetic torque; (c) DC-DC output voltage; and (d) inverter line-line voltage

In general, interfacing power converters with PMSG introduces low frequency harmonics which tends to distort output wave form of the PMSG and also heighten the level of grid THD. Multilevel inverter and line filtering are used to complement each other in combating level of grid THD. LC line filter is effective but it does have an effect of imposing limitations to the grid voltage.

CONCLUSIONS

In this paper, a topology to connect two wind turbine generators (WTGs) to the grid using only one 3-level NPC inverter has been suggested. Advanced hill climbs search (AHCS) MPPT algorithm was developed to generate proper duty cycle for the control of single stage DC/DC boost converter. In the proposed system, the outputs of the DC/DC converter of each WTG were connected in series to provide a single DC source for the high-power inverter**.** It was assumed that the two WTGs are similar in all respect so that the voltage across the DC link capacitors is twice the output voltage from any one of the WTG. The proposed system was simulated using MATLAB/SIMULINK and the results for all the parameters verified were satisfactory. The NPC inverter was controlled using SPMW technique and the output voltage was obtained to be three levels. Simulation results had also shown that WECs demonstrated variability in its parameters owing to variations in wind speeds. This introduces a new factor of ambiguity on the grid and constitutes a lot of problems to the power system grid integrity i.e. power system security, power system stability and power quality.

ACKNOWLEDGEMENT

My sincere appreciation goes to my employer, Waziri Umaru Federal Polytechnic Birnin Kebbi, Nigeria and Tertiary Education Trust Fund for their moral and financial supports. My sincere gratitude goes to Prof. Dr. Saad Mekhilef for his valuable and constructive suggestions during the planning and development of this research work.

REFERENCES

Aliprantis, D., Papathanassiou, S., Papadopoulos, M., & Kladas, A. (2000, August 28-30). Modeling and control of a variable-speed wind turbine equipped with permanent magnet synchronous generator. In *Proceeding of International Conference on Electrical Machines* (pp. 558-562). Helsinki, Finland.

Bhende, C., Mishra, S., & Malla, S. G. (2011). Permanent magnet synchronous generator-based standalone wind energy supply system. *IEEE Transactions on Sustainable Energy, 2*(4), 361-373.

Bull, S. R. (2001). Renewable energy today and tomorrow. *Proceedings of the IEEE, 89*(8), 1216-1226.

Chompoo-Inwai, C., Lee, W. J., Fuangfoo, P., Williams, M., & Liao, J. R. (2005). System impact study for the interconnection of wind generation and utility system. *IEEE Transactions on Industry Applications, 41*(1), 163-168.

Dahono, P. A., & Purwadi, A. (1995, February 21-24). An LC filter design method for single-phase PWM inverters. In *Proceedings of 1995 International Conference on Power Electronics and Drive Systems (PEDS 95)* (pp. 571-576). Singapore.

David, P. P. (2010). *Sustainable power production and transportation. assignment II*. Chalmers University of Technology, Sweden. Retrieved December 13, 2015 from https://www.google.com.my/?gws_rd=cr

,ssl&ei=YrFHVLztCczauQShhICoAQ#q=sustainable+power+production+and+transportation+by+Pei yuan+Poopak+David

Do, D. T., & Nguyen, M. K. (2018). Three-level quasi-switched boost t-type inverter: analysis, PWM control, and verification. *IEEE Transactions on Industrial Electronics, 65*(10), 8320-8329.

Eltamaly, A. M., Alolah, A., & Farh, H. M. (2013). Maximum power extraction from utility-interfaced wind turbines. In H. Arman & I. Yuksel (Eds.), *New Developments in Renewable Energy* (pp. 159-192). Rijeka, Crotia: InTech.

Guo, X., Wei, B., Zhu, T., Lu, Z., Tan, L., Sun, X., & Zhang, C. (2018). Leakage current suppression of three-phase flying capacitor PV inverter with new carrier modulation and logic function. *IEEE Transactions on Power Electronics, 33*(3), 2127-2135.

Jayasinghe, S. D. G., Vilathgamuwa, D. M., & Madawala, U. K. (2010, November 7-10). Connecting two wind turbine generators to the grid using only one three level NPC inverter. In *IECON 2010-36th Annual Conference on IEEE Industrial Electronics Society* (pp. 3263-3268). Glendale, AZ, USA.

Li, H., & Chen, Z. (2008). Overview of different wind generator systems and their comparisons. *IET Renewable Power Generation, 2*(2), 123-138.

Mohamed, K. H., Ibrahim, T. B., & Saad, N. B. (2013). Boost converter design with stable output voltage for wave energy conversion system. *International Journal of Information Technology and Electrical Engineering, 2*(1), 10-15.

Nguyen, M. K., Tran, T. T., & Lim, Y. C. (2019). A family of pwm control strategies for single-phase quasi-switched-boost inverter. *IEEE Transactions on Power Electronics, 34*(2), 1458-1469.

Nishikata, S., & Tatsuta, F. (2010). A new interconnecting method for wind turbine/generators in a wind farm and basic performances of the integrated system. *IEEE Transactions on Industrial Electronics, 57*(2), 468-475.

Peng, F. Z., Qian, W., & Cao, D. (2010, June 21-24). Recent advances in multilevel converter/inverter topologies and applications. In *The 2010 International Power Electronics Conference-ECCE ASIA* (pp. 492-501). Sapporo, Japan.

Porselvi, T., & Muthu, R. (2014). Wind energy conversion system with boost converter and CHB MLI with single DC input. *International Journal of Engineering and Technology (IJET), 6*(1), 138-145.

Reidy, A., & Watson, R. (2005, June 16). Comparison of VSC based HVDC and HVAC interconnections to a large offshore wind farm. In *IEEE Power Engineering Society General Meeting, 2005* (pp. 1-8). San Francisco, CA, USA.

Samadaei, E., Sheikholeslami, A., Gholamian, S. A., & Adabi, J. (2018). A square T-type (ST-Type) module for asymmetrical multilevel inverters. *IEEE Transactions on Power Electronics, 33*(2), 987-996.

Samanvorakij, S., & Kumkratug, P. (2013, April 6-7). Modeling and Simulation PMSG based on wind energy conversion system in MATLAB/SIMULINK. In *Proceeding of the Second International Conference on Advances in Electronics and Electrical Engineering* (pp. 37-41). Bangkok, Thailand.

Sener, A. Y. A. K. (2016, August 28-31). *Modeling and Simulating of a Wind Energy System Which is Connected to Grid with Static Transfer Switch.* Paper presented at the European Conference on Renewable Energy Systems, Istanbul, Turkey.

Veilleux, E., & Lehn, P. W. (2010). Interconnection of direct-drive wind turbines using a series-connected DC grid. *IEEE Transactions on sustainable energy, 5*(1), 139-147.

Yaramasu, V., Dekka, A., Durán, M. J., Kouro, S., & Wu, B. (2017). PMSG-based wind energy conversion systems: survey on power converters and controls. *IET Electric Power Applications, 11*(6), 956-968.

Pertanika J. Sci. & Technol. 27 (4): 2189 - 2200 (2019)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Efficacy of Silvernano Bacterial Cellulose Insole on *Micrococcus sedentarius* Growth

Kanchana Sitlaothaworn[1]* and Manussawee Dechkla[2]

[1]Department of Microbiology Industry, Faculty of Science and Technology, Suan Sunandha Rajabhat University, U-thong Nok Road, Dusit, Bangkok 10300, Thailand
[2]Department of Biology, Faculty of Science and Technology, Suan Sunandha Rajabhat University, U-thong Nok Road, Dusit, Bangkok 10300, Thailand

ABSTRACT

Silver nanoparticles (Ag^0) impregnated into bacterial cellulose (BC) by immersion in an antimicrobial concentration of 1,700 µg/ml (10 mM) $AgNO_3$ solution were used to synthesize silver nanoparticle-impregnated bacterial cellulose (Ag^0BC) composites. Ag^0BC incorporated into insertable footpad was developed for reduction of foot odor associated with pitted keratolysis, primarily caused by *Micrococcus sedentarius*. A scanning electron microscope (SEM) revealed the porous structure of Ag^0BC fibers impregnated with Ag^0 particles. The antibacterial activity of Ag^0BC for *Micrococcus sedentarius* revealed a growth inhibition zone of 30 mm in diameter by disc diffusion method. The minimal inhibitory concentration (MIC) and minimal bactericidal concentration (MBC) of $AgNO_3$ solution against *Micrococcus sedentarius* were 0.42 µg/ml and 0.83 µg/ml, respectively. The effficiency of an Ag^0BC footpad applied to shoe samples was measured via total bacterial counts over a time period of 14 days. As expected, the percentage of surviving bacteria in three different types of shoe samples was gradually decreased in the period beginning on the 2nd day. Moreover, no bacterial growth was found in any shoe sample within 8–14 days. However, the cytotoxicity of 5.31–1,700 µg/ml $AgNO_3$ concentrations revealed 50% growth inhibition (IC_{50}) on human dermal fibroblast cells by resazurin microplate assay (REMA). Thus, it can be concluded that Ag^0BC, synthesized with lower doses of $AgNO_3$ concentration (< 1,700 µg/ml), may be able to apply in Ag^0BC footpad production for inhibition of *Micrococcus sedentarius* growth and prevention of foot odour in different types of footwears.

Keywords: Bacterial cellulose, foot odour, footpad, *micrococcus sedentarius*, silver nanoparticles

ARTICLE INFO

Article history:
Received: 13 November 2018
Accepted: 13 May 2019
Published: 21 October 2019

E-mail addresses:
kanchana.si@ssru.ac.th (Kanchana Sitlaothaworn)
manussawee.de@ssru.ac.th (Manussawee Dechkla)
* Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

INTRODUCTION

Silver (Ag) is a well-known metal that can be used to inhibit the growth of both Gram-positive and -negative microorganisms (Klasen, 2000a; Klasen, 200b; Morones et al., 2005; Barud et al., 2011). The effectiveness of silver nanoparticles (Ag^0), generated on bacterial cellulose (BC), in the antimicrobial agents is due to their large surface area which can effectively interact with microbes. The inhibitory processes involve the penetration of Ag particles into the bacterial cell wall, followed by the binding of Ag ions to SH (sulfhydryl) groups of metabolic enzymes and proteins. The cytotoxic mechanism of Ag particles affects cellular metabolism which leads to inhibition of cell growth by the destruction of both the electron transport system in the respiratory chain and substantial movement across cell membranes (Cho et al., 2005). Today, silvernano technology is currently used as antibacterial additives in many health care products (Vance et al., 2015).

Bacterial cellulose (BC) was produced from *Gluconacetobacter xylinus* (*Acetobacter aceti* subsp. *xylinus* or *A. xylinus*). BC is an alternative material in the production of shoe pads due to its many advantageous properties. According to these properties, BC is different from plant cellulose due to being composed of lignin and hemicellulose as its major components. BC structure is a 3-dimensional network composing of cellulose fibers, which are relatively well absorbed and store any liquid due to its porous nature. The structure exhibits high water permeability which drains an adhesive irrigation fluid over the dry surfaces (Czaja et al., 2007). BC has been used extensively in food applications and also successfully applied to the medical field, such as for artificial skins and wound dressing (Hutchens et al., 2006; Czaja et al., 2007; Shahbazzadeh et al., 2011; Wu et al., 2014). However, BC does not have the ability to kill or inhibit the growth of microorganisms (Maneerung et al., 2008).

Micrococcus sedentarius, a Gram-positive bacterium, is a major cause of pitted keratolysis in tropical regions (Almeida Jr et al., 2016). Pitted keratolysis is caused by cutaneous infection of the stratum corneum by *Micrococcus sedentarius*, although *Dermatophilus congolensis* and *Corynebacterium* species have also been implicated. Pitted keratolysis has a worldwide distribution and commonly affects people wearing occlusive footwear for prolonged periods (Kaptanoglu et al., 2012). Under appropriate conditions such as hyperhidrosis in heat and humid environments, these bacteria produce proteolytic enzymes which digest the stratum corneum of the sole's surface (Vlahovic, 2016). Plantar hyperhidrosis predisposes to foot odour associated with pitted keratolysis and pitted lesions, affecting quality of life and emotional well-being. The treatment of pitted keratolysis is based on topical antibiotics including natural antimicrobial agents; such as Gallnut (Yamuna & Sudha, 2013), fragrant agents; such as citral, citronellal and geraniol (Ara et al., 2006) and avoiding foot humidity; however, resistant cases have been reported by diverse patients (Singh & Naik, 2005; Blaise et al., 2008; Makhecha et al., 2017).

Thus, this study aims to create a silvernano bacterial cellulose (Ag^0BC) which is applied to the footpad, for inhibition of *Micrococcus sedentarius* growth, a major factor associated with foot odour (smelly feet) in tropical countries (Sokolnicki et al., 2006). There are currently many products that claim to reduce undesirable smells; for example, deodourant sprays, antiperspirant products and copper-infused socks (Quaranta et al., 2011). However, some of these products can cause allergies, irritation to the skin and increased antimicrobial resistance (Singh & Naik, 2005). Therefore, the antibacterial efficiency of Ag^0BC incorporated into an insertable footpad on shoe samples is investigated in this research to develop the current and future of quality footpad production for preventing foot odour.

MATERIALS AND METHODS

Microorganisms and Growth Conditions

M. sedentarius (strain DMST 9365) was purchased from the Department of Medical Sciences (DMSC, Thailand). Cells were maintained on nutrient broth (NB) and transfered into the nutrient agar (NA) plates. The inoculated plates were further incubated at 37 °C for 24 hours (Piumnoppakun, 2009).

Acetobacter xylinum (strain TISTR 975) was obtained from the Microbiological Resources Centre, Thailand Institute of Scientific and Technological Research (TISTR, Thailand). Cells were grown in glucose yeast extract (GYE) broth at 30°C for 3–5 days. The culture was spread uniformly and futher incubated on GYE agar plates.

Antibacterial Activities of Silver Nitrate by Disc Diffusion Method

M. sedentarius was grown in NB until the optical density (OD_{625}) reached 0.08–0.1 (approximately of 10^8 cfu/ml) and spread onto the NA agar plate. Then, 1,700 µg/ml (10 mM), 170 µg/ml (1 mM) and 17 µg/ml (0.1 mM) of $AgNO_3$ solutions was dropped on the the disc which was placed on a cultured agar plate. After this plate was incubated at 37°C for 24 h, the diameter of inhibition zones was examined by comparing with a control disc (without $AgNO_3$ solution).

Production and Purification of Bacterial Cellulose

A single colony of *Acetobacter xylinum* grown on a GYE agar plate was transfered into 50 ml GYE broth. A BC pellicle was produced on the surface of the culture medium after incubation at 30°C for 24 h. The pellicle emerged with a diameter of 8.5 cm and a wall thickness of 0.6 cm; pellicles were harvested and purified in 1% NaOH and repeatly washed in distilled water followed by the method of Maneerung et al. (2008), before use in further assays.

Synthesis of Silvernano Bacterial Cellulose (Ag^0BC)

Ag^0BC was produced according to the Maneerung et al. (2008) protocol. Briefly, silver nanoparticles were impregnated into a BC pellicle by immersing in 1,700 µg/ml $AgNO_3$ solution for one hour. The BC was then washed with ethanol for 30 sec, followed by soaking in sodium boro-hydride ($NaBH_4$) solution at a proportion of 1 $AgNO_3$: 100 $NaBH_4$ for 10 min. Ag^0BC were rinsed in ultra-pure water, frozen at -20^{o}C and freeze-dried before use in the following steps.

Antimicrobial Activity of Silvernano Bacterial Cellulose

Freeze-dried Ag^0BC was cut into disc-shapes of 6 and 15 mm and placed on an *M. sedentarius*-culture agar plate ($1x10^8$ cfu/ml). The plates were incubated at 37^{o}C for 24 hours and the diameter of inhibition zone was measured by comparing with pure BC.

Antimicrobial Activity of Ag^0BC Footpad

Ag^0BC was cut into a foot-shape and placed on three different types of shoe samples: A, B and C. The bacterial cells were collected with sterile cotton swabs on the shoe-surface interface (2x2 inch^2). Each swab sample were placed and diluted in 10 ml of 0.85% NaCl solution. The bacterial suspensions were transferred to agar plates and incubated at 37^{o}C for 24 h. Then, the total number of surviving cells were monitored by counting the number of colony-forming units (CFUs) every 2 days for 14 days, comparing with a control shoe (without Ag^0BC footpad).

Minimal Inhibitory Concentration (MIC) and Minimum Bactericidal Concentration (MBC) Assay

Stock solutions of 100 mM $AgNO_3$ were prepared by dissolving 17,000 µg of $AgNO_3$ in 1 ml (17 mg/ml) of distilled water. The solutions were diluted in NB to obtain an $AgNO_3$ concentration of 3,400 µg/ml. Then, serial dilutions of $AgNO_3$ solution were prepared in NB with concentrations ranging between 0.0033 and 1,700 µg/ml (20 dilutions). Fifty µl of each dilution was added to 50 µl of bacterial suspension in a 96-well microtitre plate. The plates were incubated at 37^{o}C for 24 hours and examined for bacterial growth by comparing to the controls. MIC was defined as the lowest $AgNO_3$ concentration for inhibiting visible growth of *M. sedentarius*. The macroscopically clear wells were selected as the MIC value. MBC was determined by transferring culture from the wells observing no visible growth to NA plates. The plates were incubated at 37^{o}C for 24 hours. The lowest $AgNO_3$ concentration showing no visible growth on NA plates was considered to be the MBC value. The experiment for each dilution was performed in triplicate and was compared with control wells, culture medium for sterility testing (well 21) and culture medium mixed with microorganism for bacterial growth (well 22).

SEM Characterization of Ag^0BC

Scanning electron microscopy was used to observe the formation of silver nanoparticles on BC pellicles before and after exposure to $AgNO_3$. The surface images of freeze-dried BC membranes were observed and taken by SEM (JEOL model JSM-5410LV, Scientific and technological research equipment centre [STREC], Chulalongkorn University, Thailand).

Cytoxicity Assay

A Resazurin Microplate Assay (REMA) was used to evaluate the toxicity effect of $AgNO_3$ solution on human dermal fibroblast, neonatal (HDFn, C-004-5C). The assay was performed by the National Center for Genetic Engineering and Biotechnology (BIOTEC, Thailand), according to the method of O′Brien et al. (2000). Briefly, 10^4 cells were treated with various $AgNO_3$ concentrations, followed by addition of 10% Resazurin solution and measurement of fluorescence signals. Cell viability and 50% inhibitory concentration (IC_{50}) values were calculated using SOFTMax Pro software (Molecular Devices, USA). Ellipticine and 1% DMSO were used as a positive and a negative control, respectively.

RESULTS AND DISCUSSIONS

Determination of $AgNO_3$ Concentration in *M. sedentarius* Growth Inhibition

Antibacterial activity of three different $AgNO_3$ concentrations (1,700, 170 and 17 µg/ml) against *M. sedentarius* was determined by measuring the diameter of inhibition zones. It was found that the $AgNO_3$ solution exhibited zones of inhibition at least 15 mm in diameter. In the presence of 1,700 and 170 µg/ml $AgNO_3$ solution, the diameter of growth inhibition rings against *M. sedentarius* was 21 and 15 mm, respectively. In constrast, no inhibition zone was observed with 17 µg/ml $AgNO_3$ solution (Figure 1). The results demonstrated that $AgNO_3$ solution has the ability to inhibit the growth of *M. sedentarius*. The concentration of 1,700 µg/ml $AgNO_3$ solution was used for production of Ag^0BC in the following step.

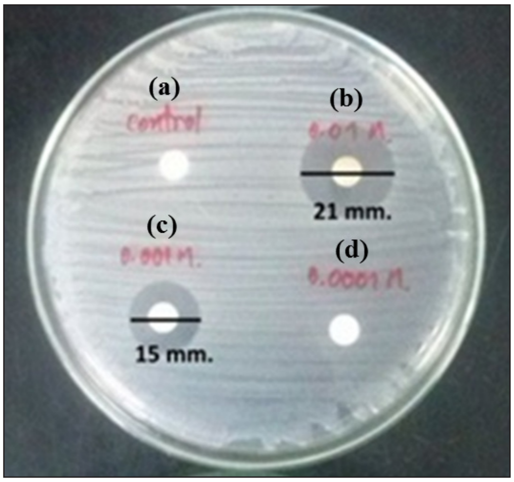
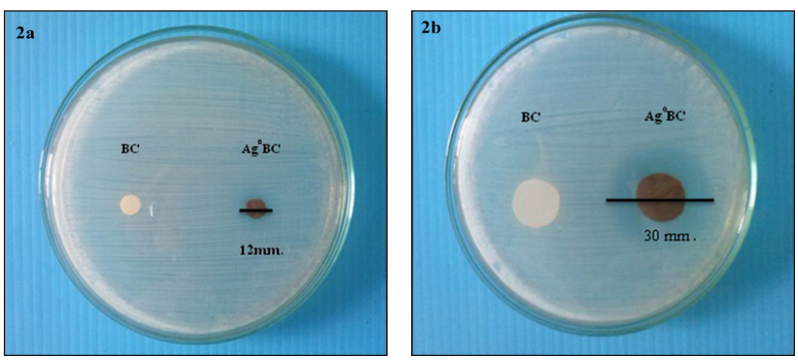


Figure 1. Three different concentrations of $AgNO_3$ solution in inhibiting *M. sedentarius* growth. Agar plate shows the zones of inhibition generated by (b) 1,700 µg/ml [0.01 M] and (c) 170 µg/ml [0.001 M] $AgNO_3$ solution, comparing with (a) control (without $AgNO_3$ solution) and (d) 17 µg/ml [0.0001 M]. Long straight lines indicate the diameter of the zones of inhibition.

Antimicrobial Activity of Silvernano Bacterial Cellulose in the Inhibition of *M. sedentarius* Growth

Freeze-dried Ag^0BC pellicles, with diameters of 6 and 15 mm, were placed on *M. sedentarius* agar-plate culture for determination of bacterial growth inhibition. The diameter of inhibition zone on Ag^0BC pellicles was 12 and 30 mm around 6 and 15 mm discs, respectively (Figure 2), comparing with the BC without Ag^0. However, the zone of inhibition exhibited on Ag^0BC around the 6 mm disc was smaller than that on the agar diffusion test around the 21 mm disc, with the same concentration of 1,700 µg/ml $AgNO_3$. These differences may result from the distribution of Ag^0 particles which are not thoroughly absorbed in BC pellicle pores during the Ag^0BC production process. Yang et al. (2012) has reported that the most typical carbon source in fermentation processes affected the structure of the BC pellicle; such as surface, pore size and shape. Thus, the BC structural changes have an effect on the amount of Ag^0 particle size distribution in BC pellicles during the Ag^0BC production process. However, the results indicated that the Ag^0BC was able to inhibit the growth of *M. sedentarius* in the presence of Ag^0 particles, while the individual BC alone expressed an undetectable antimicrobial activity.



(a) (b)

Figure 2. Antimicrobial activity of Ag^0BC in 6 mm (a) and 15 mm (b) diameter discs against *M. sedentarius*, comparing with pure BC. Black straight-lines indicate inhibition zone diameters of 12 and 30 mm around 6 and 15 mm Ag^0BC pellicles, respectively.

Efficiency of Silvernano Bacterial Cellulose Footpad for Inhibition of Bacterial Growth in the Shoe Samples

Foot-shaped Ag^0BC was placed in three different types of shoe samples: A, B and C. The inhibition of bacterial growth among all shoe samples was determined by counting the number of total bacteria every 2 days for 14 days, comparing with the control shoe (without Ag^0BC footpad). After the Ag^0BC footpad was applied to shoe samples for two days, the percentage of bacterial survival was shown to rapidly decrease to 14.7%, 0.82% and 1.67% from A, B and C shoe samples, respectively, comparing with the control shoe (84.37%)

(Figure 3). Moreover, no viable cells were found in the shoe samples A, B and C from 8 to 14 days, comparing with the control shoe (84.37%). These results indicated that the Ag^0BC footpad has additional effective inhibition of bacterial growth in all shoe samples.

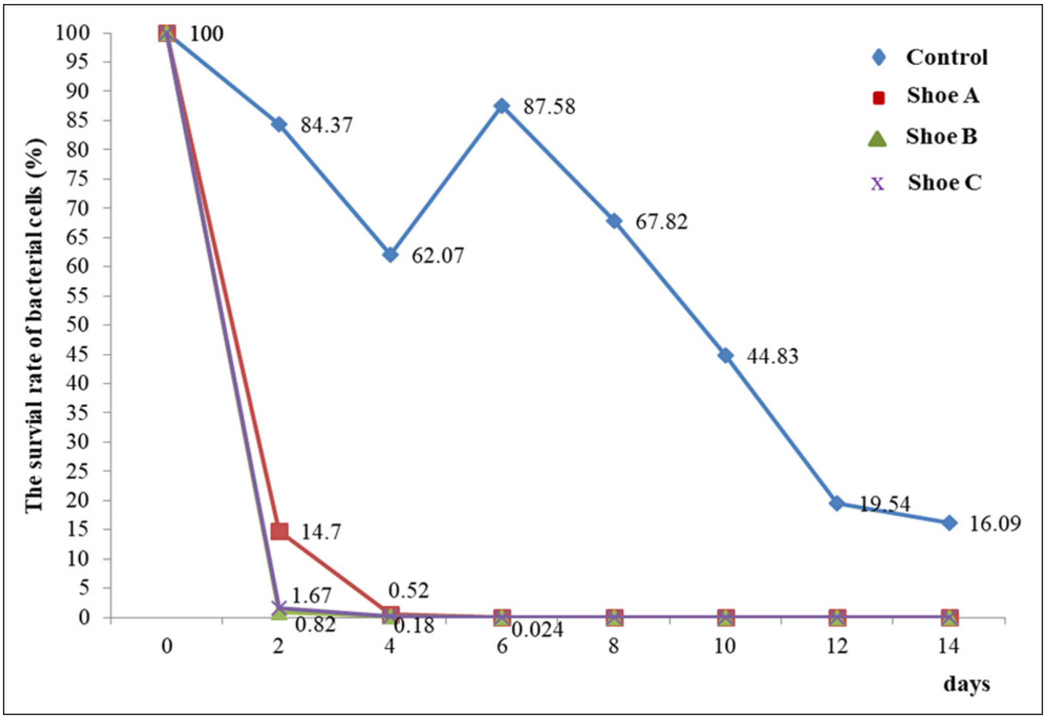


Figure 3. Percentage of bacterial survival in three types of shoe samples: A, B and C, supplied with Ag^0BC footpads for 0–14 days. From 8 to 14 days, the results demonstrated a 0% cell survival rate from all shoe samples, A(■) B(▲) and C(x), compared with 67.82–16.09% of cell survival rates in control shoes (♦).

Measurement of Minimal Inhibitory Concentration (MIC) and Minimum Bactericidal Concentration (MBC)

Different concentrations of $AgNO_3$ solution, 0.0033–1,700 µg/ml, were used to determine the growth inhibition effect on *M. sedentarius* by MIC and MBC evaluations. From the MIC result, the growth of *M. sedentarius* was found to completely inhibit in the number (No.) of wells 1 through 13 (Table 1), comparing with the wells of controls (wells No.21 and 22). Therefore, the MIC of $AgNO_3$ solution against *M. sedentarius* was 0.42 µg/ml (well No.13). However, $AgNO_3$ concentration at the MIC level (0.42 µg/ml) did not completely inhibit the growth of *M. sedentarius* on NA medium (Figure 4, sector 5, well No.13). The concentration of 0.83 µg/ml $AgNO_3$ solution (Figure 4, sector 4, well No.12) significantly prevented growth of *M. sedentarius* on agar medium. Therefore, for the MBC result, the lowest concentration of $AgNO_3$ solution that inhibited the visible growth of *M. sedentarius* was 0.83 µg/ml.

Table 1
Effect of silver nitrate concentrations on the growth of M. sedentarius *at the minimum inhibitory concentration (MIC) level*

Number of Wells	$AgNO_3$ Concentrations (μg/ml)	*M. sedentarius* Growth*
1-13	1,700-0.42	-
14	0.21	+
15	0.10	+
16	0.05	+
17	0.026	+
18	0.013	+
19	0.0065	+
20	0.0033	+
21	Agar	-
22	*M. sedentarius* culture	+

* Plus (+) and minus (-) symbols define as the presence and absence of *M. sedentarius* growth, respectively

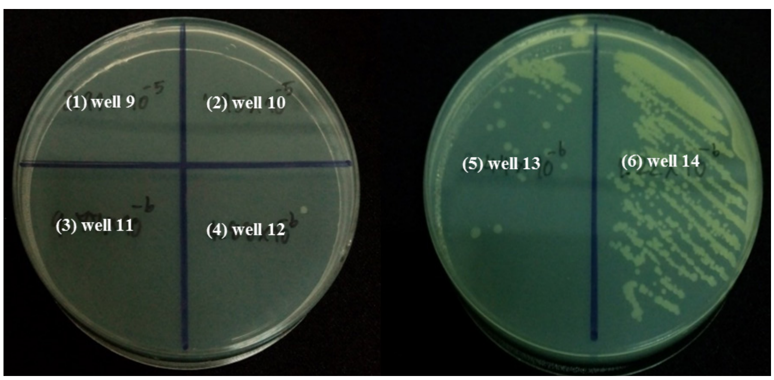
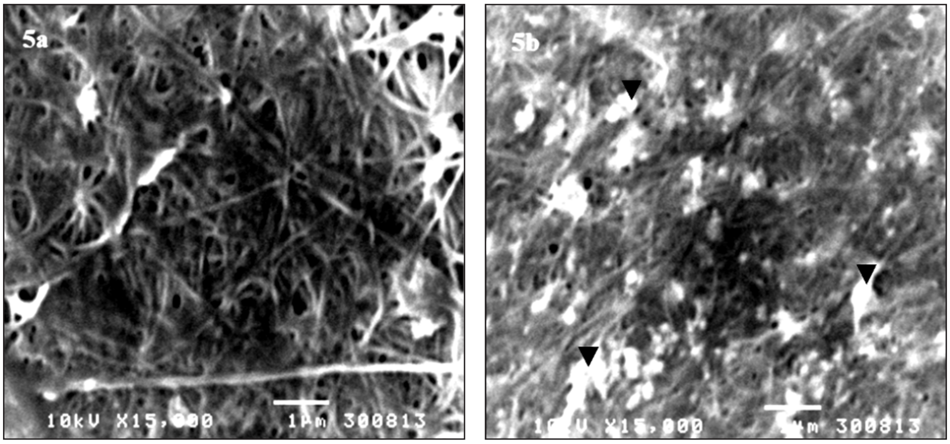


Figure 4. MBC level of different silver nitrate concentrations that inhibited the visible growth of *M. sedentarius* is shown in sector 4 (well No.12) of agar plate

Adsorption Efficiency of Silver Nanoparticles into Bacterial Cellulose by Scanning Electron Microscopy (SEM)

The 3-dimensional structure of Ag^0BC illustrated the impregnation of Ag^0 on BC surface pellicles and within the BC porous network by SEM (Figure 5). In contrast, the natural fibers of the BC structure demonstrated no impregnation of Ag^0 particles into BC networks. However, the distribution of silver nanoparticles on Ag^0BC pellicles was diffused with an irregular pattern, resulting in a size of Ag^0BC inhibition zone of less than the drops of silver solution on the disc. Therefore, the method of BC pellicle formation is necessary to develop and rearrange the fibrous network of the BC structure; such as the number of BC pores and the average size of BC pores that are suitable for impregnation and release of Ag^0 through BC pellicles (Kvitek et al., 2008).



(a) (b)

Figure 5. SEM of (a) naturally occurring BC pellicle and (b) impregnated BC pellicle with silver nanoparticles as indicated by the black arrows. The scale bar represents 1 µm.

Cytoxicity of Silver Nitrate Solution on Human Neonatal Dermal Fibroblasts with Resazurin Microplate Assay (REMA)

The cytotoxicity of $AgNO_3$ solution upon human neonatal dermal fibroblasts (HDFn, C-004-5C) was determined with the REMA plate method. By using the final concentration of $AgNO_3$ solution over the range of 5.31 to 1,700 µg/ml, IC_{50} values for dermal fibroblast cells were found to be 13.99 ± 2.50 µg/ml (Table 2). These results correlated with the study of nanosilver (Nano-Ag) effects against two normal human cell lines, demonstrating IC_{50} values of 6.33 and 6.68 µg/ml in mesenchymal stem cell and fibroblast HF2, respectively, with the MTT assay (Shahbazzadeh et al., 2011), suggesting the concentration-dependent toxicity effect of silver nanoparticles in human cell viability *in vitro*. However, the concentration of 1,700 µg/ml $AgNO_3$ solution selected for Ag^0BC production was toxic to the dermal cells. Therefore, the appropriately non-toxic $AgNO_3$ concentration implicated in silver nanoparticle synthesis should be less than 1,700 µg/ml for Ag^0BC footpad production.

Table 2
The IC_{50} values of silver nitrate solution on primary neonatal human dermal fibroblast (HDFn)

Sample	Final Concentrations (µg/ml)	IC_{50} ± SD (µg/ml)	Toxicity on HDFn
Cell + DMSO (Negative Control 1)	1% DMSO	-	non-cytotoxic
Cell + Distilled water (Negative Control 2)	Distilled water	-	non-cytotoxic
Ellipticine (Positive Control)	0.31-10.00	3.31 ± 0.75	cytotoxic
$AgNO_3$ solution	**5.31-1,700**	**13.99 ± 2.50**	cytotoxic

CONCLUSIONS

The present study revealed the 3-dimensional Ag^0BC structure that exhibited fibrous networks containing the impregnation of silver nanoparticles in porous BC surfaces. Ag^0BC pellicles generated from 1,700 µg/ml (10 mM) $AgNO_3$ concentration were able to completely inhibit the growth of *M. sedentarius* according to the disc diffusion method. Moreover, Ag^0BC incorporated into an insertable footpad was effective in reducing the total number of live bacteria in three different types of shoe samples from 8 to 14 days. However, further Ag^0BC analyses need to be conducted to find an appropriate $AgNO_3$ concentration for production of Ag^0BC-applied footpad which is less toxic to dermal cells (Farag et al., 2015) and eliminates bacterial growth in various types of footwears.

ACKNOWLEDGMENT

This study was supported by Suan Sunandha Rajabhat University.

REFERENCES

Almeida Jr, H. L. D., Siqueira, R. N., Meireles, R. D. S., Rampon, G., & Castro, L. A. S. D. (2016). Pitted keratolysis. *Anais Brasileiros de Dermatologia, 91*(1), 106-108.

Ara, K., Hama, M., Akiba, S., Koike, K., Okisaka, K., Hagura, T., … & Tomita, F. (2006). Foot odor due to microbial metabolism and its control. *Canadian Journal of Microbiology*, *52*(4), 357-364.

Barud, H. S., Regiani, T., Marques, R. F. C., Lustri, W. R., Messaddeq, Y., & Ribeiro, S. J. L. (2011). Antimicrobial bacterial cellulose-silver nanoparticles composite membranes. *Journal of Nanomaterials*, *8*(6), 1-8.

Blaise, G., Nikkels, A. F., Hermanns-Le, T., Nikkels-Tassoudji, N., & Pierard, G. E. (2008). Corynebacterium-associated skin infections. *International Journal of Dermatology*, *47*(9), 884-890.

Cho, K. H., Park, J. E., Osaka, T., & Park, S. G. (2005). The study of antimicrobial activity and preservative effects of nanosilver ingredient. *Electrochimica Acta*, *51*, 956-960.

Czaja, W. K., Young, D. J., Kawecki, M., & Brown, R. M. (2007). The future prospects of microbial cellulose in biomedical applications. *Biomacromolecules*, *8*(1), 1-12.

Farag, S., Ibrahim, H. M., Asker, M. S., Amr, A., & El-Shafaee, A. (2015). Impregnation of silver nanoparticles into bacterial cellulose: Green synthesis and cytotoxicity. *International Journal of ChemTech Research, 8*(12), 651-661.

Hutchens, S. A., Benson, R. S., Evans, B. R., O′Neill, H. M., & Rawn, C. J. (2006). Biomimetic synthesis of calcium-deficient hydroxyapatite in a natural hydrogel. *Biomaterials*, *27*(26), 4661-4670.

Kaptanoglu, A. F., Yuksel, O., & Ozyurt, S. (2012). Plantar pitted keratolysis: a study from non-risk groups. *Dermatology reports, 4*(e4), 14-16.

Klasen, H. J. (2000a). A historical review of the use of silver in the treatment of burns. II. Renewed interest for silver. *Burns*, *26*(2), 131-138.

Klasen, H. J. (2000b). Historical review of the use of silver in the treatment of burns. I. Early uses. *Burns*, *26*(2), 117-130.

Kvitek, L., Panacek, A., Soukupova, J., Kolar, M., Vecerova, R., Prucek, R., … & Zboril, R. (2008). Effect of surfactants and polymers on stability and antibacterial activity of silver nanoparticles (NPs). *The Journal of Physical Chemistry C*, *112*(15), 5825-5834.

Makhecha, M., Dass, S., Singh, T., Gandhi, R., Yadav, T., & Rathod, D. (2017). Pitted keratolysis - a study of various clinical manifestations. *International Journal of Dermatology*, *56*(11), 1154-1160.

Maneerung, T., Tokura, S., & Rujiravanit, R. (2008). Impregnation of silver nanoparticles into bacterial cellulose for antimicrobial wound dressing. *Carbohydrate Polymers*, *72*(1), 43-51.

Morones, J. R., Elechiguerra, J. L., Camacho, A., Holt, K., Kouri, J. B., Ramirez, J. T., & Yacaman, M. J. (2005). The bactericidal effect of silver nanoparticles. *Nanotechnology*, *16*(10), 2346-2353.

O′Brien, J., Wilson, I., Orton, T., & Pognan, F. (2000). Investigation of the Alamar Blue (resazurin) fluorescent dye for the assessment of mammalian cell cytotoxicity. *European Journal of Biochemistry*, *267*, 5421-5426.

Piumnoppakun, N. (2009). *Development of foot spray product using essential oil in inhibitory Micrococcus sedentarius growth.* Retrieved April 16, 2009, from http://doi.nrct.go.th/ListDoi/Download/18671/be44b38a07bd95a3b67bb30c4ba6eea5?Resolve_DOI=10.14457/KU.the.2009.164

Quaranta, D., Krans, T., Santo, C. E., Elowsky, C. G., Domaille, D. W., Chang, C. J., & Grass, G. (2011). Mechanisms of contact-mediated killing of yeast cell on dry metallic copper surfaces. *Applied and Environmental Microbiology*, *77*(2), 416-426.

Shahbazzadeh, D., Ahari, H., Motalebi, A. A., Anvar, A. A., Moaddab, S., Asadi, T., … & Rahman-Nya, J. (2011). *In vitro* effect of Nanosilver toxicity on fibroblast and mesenchymal stem cell lines. *Iranian Journal of Fisheries Sciences*, *10*(3), 487-496.

Singh, G., & Naik, C.L. (2005). Pitted keratolysis. *Indian Journal of Dermatology, Venereology, and Leprology*, *71*(3), 213-215.

Sokolnicki, A. M., Fisher, R. J., Harrah, T. P., & Kaplan, D. L. (2006). Permeability of bacterial cellulose membranes. *Journal of Membrane Science*, *272*(1-2), 15-27.

Vance, M. E., Kuiken, T., Vejerano, E. P., McGinnis, S. P., Hochella Jr, M. F., Rejeski, D., & Hull, M. S. (2015). Nanotechnology in the real world: Redeveloping the nanomaterial consumer products inventory. *Beilstein Journal of Nanotechnology*, *6*(1), 1769-1780.

Vlahovic, T. C. (2016). Plantar hyperhidrosis: an overview. *Clinics in Podiatric Medicine and Surgery*, *33*(3), 441-451.

Wu, J., Zheng, Y., Wen, X., Lin, Q., Chen, X., & Wu, Z. (2014). Silver nanoparticle/bacterial cellulose gel membranes for antibacterial wound dressing: investigation *in vitro* and *in vivo*. *Biomedical Materials*, *9*(3), 1-12.

Yamuna, V., & Sudha, S. (2013). Antimicrobial activity of fabrics treated with *Quercus Infectoria* extract for foot odour control. *BioTechnology: An Indian Journal, 7*(6), 231-235.

Yang, G., Xie, J., Feng, H., Cao, Z., & Yang, X. (2012). Antimicrobial activity of silver nanoparticle impregnated bacterial cellulose membrane: Effect of fermentation carbon sources of bacterial cellulose. *Carbohydrate Polymers*, *87*(1), 839-845.

Pertanika J. Sci. & Technol. 27 (4): 2201 - 2220 (2019)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Elemental Concentration in Three Different Fish Species Captured from Oluwa River, Okitipupa, Ondo State, Nigeria

Thompson Faraday Ediagbonya* and Ramond Babatunde Adesokan

Department of Chemical Sciences, Faculty of Science, Ondo State University of Science and Technology, Okitipupa, Nigeria

ABSTRACT

There has been a rapid turnaround from meat consumption to fish consumption in Nigeria over the last few years. As at 2013, the rate of fish consumption in Nigeria has risen to 2.6 million tonnes with a per capita consumption of 13.5 kg. Heavy metals bio-accumulate in fishes especially when the water body is polluted. Hence, humans are at risk of being affected by these metals via the consumption of contaminated fish, fish products and other aquatic foods captured from a contaminated river for this reason, there is need to monitor the level of heavy metal concentration in water bodies where they are caught. Analyses were done on fourteen (14) elements (Fe, Ni, Cr ,Cu , Mn, As, Zn,Ca, Ti, Se,Rb,K, Sr, Co) in three different fish species were analysed using Energy Dispersive X-ray florescence (EDXRF). The result showed that the mean values of Cr, Mn, Fe, Ni, Cu, Zn and As were all Higher than the recommended FAO/WHO standards. Generally, metal concentration in the gills is higher than that of the muscles for each species except Zn, Se and Fe in which is higher in the muscle for Tilapia fish and Fe, K and Rb which is higher in the muscle for Catfish. The target hazard quotient (THQ) of each metal through consumption of fishes from Oluwa river for both adults and children increased in the following order: Fe < Ni < Cr < Cu < Mn < As.

Keywords: Element, hazard index, target hazard quotient, X-Ray fluorescence

ARTICLE INFO

Article history:
Received: 10 December 2018
Accepted: 25 June 2019
Published: 21 October 2019

E-mail addresses:
tf.ediagbonya@gmail.com/tf.ediagbonya@osustech.edu.ng (Thompson Faraday Ediagbonya)
adesokan@gmail.com (Ramond Babatunde Adesokan)
* Corresponding author

INTRODUCTION

Elements can be grouped as metal, non-metal or metalloid or essential based on their properties. When a metal is heavy, it has relatively high density, essential, it is harmless, metallic, conduct heat, toxic and it has larger amounts or certain form (Fergusson, 1990). Heavy metals are those

ISSN: 0128-7680
e-ISSN: 2231-8526

metallic elements with relative atomic masses higher than iron (Duffus, 2002). Elements enter the water body either naturally or through anthropogenic means. They are inevitably a natural constituent of the human environment. In riverine areas, metal pollution sources from direct atmospheric deposition, geologic weathering or through industrial and agricultural waste discharge (Dawson & Macklin, 1998; Kakulu et al., 1987). Water bodies are one of the receiving ends for pollutants, most especially heavy metals. These metals are deposited in the sediments of the benthic layer of the river or exist as ions in water. These metals are then taken in by aquatic organisms such as planktons, fishes and other invertebrates where they bio-accumulate and are transferred through the food chain to man, where it biomagnifies from one tropic level to another (Edward et al., 2013; Oti-Wilberforce et al., 2016). There are four routes through which metals can find their way into the body of fishes and they are; penetration through the skin, through the gills via respiration, through the sediments which they feed on, and through the intake of water (Odeemelam et al., 1999).

There has been a rapid turnaround from meat consumption to fish consumption in Nigeria over the last few years. As at 2013, the rate of fish consumption in Nigeria has risen to 2.6 million tonnes with a per capita consumption of 13.5 kg (Maureen, 2013). This is because fishes are a richer source of protein and in addition to that, it also provides omega-3 fatty acid which is known to prevent heart problem (Narain & Nunes, 2007). However, pollution of rivers by metallic elements is a great threat to consumer fishery products (Terra et al., 2008). This is because despite its nutritional value, consumption of fish contaminated with heavy metals is very hazardous to the human consumers. It has been reported that prolonged consumption of unsafe concentrations of heavy metals through foodstuff may lead to the chronic accumulations of the metals in the kidney and liver of humans causing disturbance in of numerous biochemical processes, leading to cardiovascular, nervous, kidney and bone diseases (Jarup, 2003; Trichopoulos, 1997), as heavy metals bio accumulate. Hence, humans are at risk of being affected by these metals via the consumption of contaminated fish, fish products and other aquatic foods captured from a contaminated river (Aderinola et al., 2009), for this reason, there is need to monitor the level of heavy metal concentration in water bodies where they are caught from so that for contaminated sites, the potential damage in the exposed biota can be determined so as to elucidate and solve many of the challenges in ecotoxicology, and for undisturbed sites or sites with moderate level of contamination, they can be kept under check (Awadesh, 2004).

These metallic elements can be categorised as essential e.g copper zinc, and selenium. semi-essential e.g nickel, vanadium and cobalt, and potentially toxic metals e.g aluminium, arsenic, lead, cadmium, antimony, and mercury (Szentmihalyi & Then, 2007). Even though some of these metals are essential but they can also be detrimental in highly concentrated amounts (Tüzen, 2003).

Okitipupa, the area chose (Oluwa River) for this study is a very important region with respect to the aquatic environment. Though they have no major industries but due to the fact that it is a small and developing town, they lack adequate refuse and sewage management. These waste products including effluents from minor industries most especially palm oil mills are being discharged directly into their water bodies. These pollutants may contain heavy metals that can endanger both aquatic and human life. Therefore, as a town with over 230,000 occupants according to 2006 census and a high rate of aquatic food consumption especially fishes from rivers, there is a need to find out and know the risks involved in consuming the fishes captured from this water body.

The objective of this study is to determine the elements present in the gills and muscles of three fish's species captured from Oluwa River in Okitipupa using Energy Dispersive X-ray florescence (EDXRF), compare their values with World Health Organisation (WHO) standard and to determine the target hazard quotient (THQ) and hazard index associated with consuming fishes from Oluwa River.

MATERIALS AND METHOD

The selected fish samples were from Oluwa River located in Okitipupa, Ondo state.

Oluwa River

Oluwa River is one of the major rivers in Okitipupa, Ondo state, Nigeria. Its geographical co-ordinates are latitude 4^{o}45'13.43" and longitude 6^{o}38'45.94". Aside fishes brought from neighbouring communities like Igbokoda, Igbonla, Mahin, Araromi, it is the major Source of fishes consumed in Okitipupa. Major activities in the river include fishing, transportation and sand mining. Some fish species that inhabit the river include Carps, Nile perch, Catfish, Snake head, and Tilapia.Oluwa River is linked to Irele and Igbokoda hence it serves as transport routes to these places. Transportation medium is usually by boats. Figure 1 shows the map of Oluwa River and the three locations in which the fishes were captured

Sampling

Twelve commercial fish species were caught using nets, four each for different species of fish namely; *Tilapia zilli, Heterobranchus bidorsalis and Cyprinus carpio* which were freshly harvested from the river were bought from the fishermen at the river side at about 11am and taken immediately to the laboratory for processing.

Sample Preparation

The fish samples were washed thoroughly using distilled water and dissected with clean stainless steel instruments to separate the muscle and the gills respectively according to

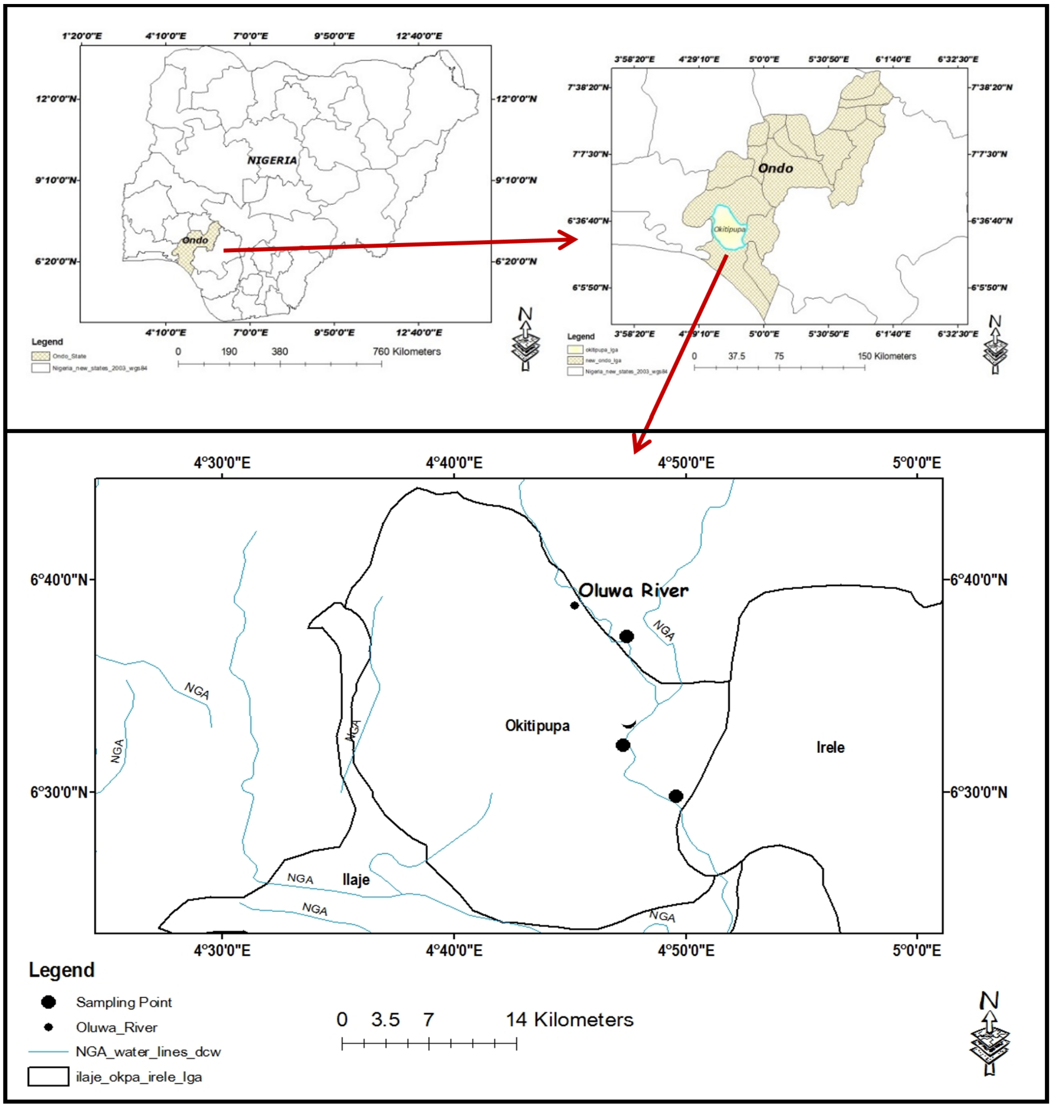


Figure 1. Shows the map of Oluwa River

FAO method (Dybem, 1993). The gills and muscles from each of the three species were rewashed thoroughly using clean water, then deionised water and kept in separate stainless steel plates. The samples were dried in an oven at 105°C until a constant weight was reached and a blender was used to reduce each dry sample to a powdered form.

Preparation and Measurement for X-Ray Fluorescence Analysis

The elemental analysis of the dried granulated fish samples was performed using the Energy Dispersive X-ray Fluorescence (EDXRF) spectrophotometer. The granulated fish sample was put in a pellet and inserted directly into the instrument. The sample was irradiated with a beam of X-rays. This primary radiation interacts with the elements in the sample to produce vacancies in the inner atomic shells, which then de-excite to produce characteristic secondary X-ray radiation. The wavelengths detected indicate which types of elements are present, and the quantity was determined from the intensity of the X-rays at each characteristic wavelength (USEPA, 1999). Each granulated fish samples was irradiated for 1000 seconds at fixed condition of 25 kV and 50 μA. The X-Ray Detector is a Model XR-100CR, high-performance thermoelectrically cooled Si-PIN photodiode, with a preamplifier. The analysis was carried out at Obafemi Awolowo University Ile-Ife.

Quality Control

Quality Assurance: Accuracy and precision were verified by using reference material (IAEA) provided by the International Atomic Agency (IAEA). Analytical results of the quality control samples indicated a satisfactory performance of heavy metal determination within the range of certified values 95–101% recovery for the metals studied.

Health Risk Assessment Parameters

Target Hazard Quotient (THQ). Potential health risk assessments based on estimated daily intake (EDI) values and target Hazard Quotient (THQ) indicated that the intakes of metals by consuming these fish species do not result in an appreciable hazard risk for the human body. To estimate the potential risk for human health derived from ingesting contaminated fish as given by FAO/WHO (2010) and the target hazard quotient (THQ) provided in the USEPA Region III Risk-based concentration table (USEPA, 2015). The target hazardous quotient (THQ) represents a complex parameter which is introduced by the US Environmental Protection Agency. It is used commonly for the assessment of the potential of non-carcinogenic risks associated with long term exposure to contaminants, such as heavy metals from food such as fish and water. Non-carcinogenic risk estimation of heavy metals consumption was determined using THQ values. THQ is a ratio of the determined dose of a pollutant to a reference level considered harmful. THQ values were determined based on the following formula (Singh et al., 2010):

$$\text{THQ} = \frac{\text{Efr} \times \text{ED} \times \text{FIR} \times C}{\text{RfDo} \times \text{B}_{\text{average wt}} \times \text{ATn} \times 10^{-3}} \quad (1)$$

where Efr is exposure frequency assumed to be 365days year^{-1}, ED is exposure duration in 70 years equivalent to an average lifetime, FIR is average daily consumption taken as 1.95×10^{-2} kg person^{-1}day^{-1}, C is concentration of metal in food sample in mg/kg, RfDo is reference dose in mg/kg day^{-1} which varies for different heavy metals, and ATn is average exposure time for non-carcinogens and is taken as 25550 days.

Hazard Index (HI). The chronic hazard index (HI) is the sum of more than one hazard quotient for multiple toxicants or multiple exposure pathways; it was calculated using the equation below (USEPA, 2011):

$$HI = \Sigma THQ. \quad (2)$$

$$HI = THQ(Ar) + THQ(Zn) + THQ(Cu) + THQ(Mn) + THQ(Ni) + THQ(Cr). \quad (3)$$

Statistical Analysis

Descriptive statistics such as mean, Standard deviation, minimum and maximum were used in mean comparison. Inferential statistical tool employed were the Pearson correlation which was used to perform the inter-relationship between the metals, while the Independent samples t-test was used to perform mean comparison between species and body parts. To explore the distribution of the physicochemical parameters, the study employed the use of Boxplot. The level of significance was set at $p<0.05$. The data analyses were carried out using the different routines in STATA version 14.

RESULTS AND DISCUSSION

Table 1 shows the mean±SD concentrations of heavy metals in Tilapia, Catfish and Carp samples. The metal with the highest concentration was calcium having its lowest value of 1115.0 mg/kg in Carp and maximum concentration value of 2650 mg/kg in Catfish. Copper had the second highest concentration having its least value of 368 mg/kg in Catfish and maximum value of 492.3 mg/kg in Carp. Zinc had the third highest concentration in the samples with a minimum value of 298.1 mg/kg and maximum value of 360.7 mg/kg in Catfish. Manganese had the fourth highest concentration with minimum value of 166.3 mg/kg in Catfish and maximum value of 244.4 mg/kg in Tilapia. Titanium had the fifth highest concentration with minimum value of 14.3 mg/kg in Catfish and maximum concentration of 25.6 mg/kg in Tilapia.

Selenium had the sixth highest concentration with minimum value 11.5 mg/kg in Carp of and maximum value of 23.8mg/kg in Tilapia. Cobalt had the seventh highest concentration with minimum value of 6.5 mg/kg in Catfish and maximum concentration of 13.1 mg/kg in Tilapia.

The level of Co detected in the fish species ranged from 6.5 – 13.1 mg/kg with the lowest concentration in Catfish and highest concentration in Tilapia. The level of Arsenic detected in the fish species ranged from 3.1 to 7.4 mg/kg, with minimum value detected in Catfish and maximum value in Tilapia. Rubidium concentration level in fish species ranged from 5.1 – 8.6 mg/kg with highest concentration in Carp and lowest level in Tilapia. Strontium level in fish species ranged from 4.1 – 9.2 mg/kg with lowest and highest concentrations in Tilapia respectively.

Table 1
Elemental concentration in Tilapia, Catfish and Carp samples collected from Oluwa River {mg/kg)

	Tilapia (n=4)			Catfish (n=4)			Carp (n=4)			FAO/WHO (1989	USFDA (1993)
	Min	Max	Mean	Min	Max	Mean	Min	Max	Mean		
K	1100	1500	1190±200	1300	1500	13700±60	2000	2000	2020±0.00	--	---
Ca	1375	2459.2	1835.68±515.15	2152	2620	2387.50±251.45	1115	1119	1117.00±2.00	--	--
Ti	17.5	25.6	20.65±3.48	14.3	20.7	17.35±3.11	14.6	14.9	14.73±0.15	---	12-13
Cr	4.1	6.40	5.10±0.96	3.1	4.4	3.87±0.50	4.4	4.9	4.67±0.25	1	-
Mn	203.7	244.40	217.61±16.46	166.3	199	181.73±16.44	215.3	215.8	215.57±0.25	1	
Fe	1200	1700	1450±0.22	1500	1700	1600±0.14	1200	1200	1240±0.01	100	--
Co	8.5	13.1	10.40±2.13	6.50	9.60	8.00±1.47	12.1	12.9	12.50±0.40	--	--
Ni	3.3	5.3	4.05±0.70	2.4	3.8	2.98±0.66	3.8	3.9	3.89±0.06	-	70-80
Cu	396.8	491.5	442.50±49.62	368	453	410±44.94	491.4	492.3	491.77±0.51	30	---
Zn	311.5	360.7	338.42±24.31	298.9	341.4	320.33±22.40	491.4	492.3	491.77±0.51	100	---
As	4.2	7.4	5.82±1.56	3.1	5.4	4.32±1.06	6.4	6.8	6.63±0.25	---	----
Se	16.0	23.8	20.58±3.34	11.5	16.6	14.07±2.60	20.4	20.9	20.67±0.25	---	---
Sr	4.1	9.2	6.57±2.53	4.4	9.1	6.58±2.16	8.1	8.5	8.27±0.21	----	---
Rb	5.1	8.2	6.53±1.44	6.4	6.9	6.60±0.20	8.1	8.6	8.33±0.25	---	----

The concentration of Nickel in the fish species ranged from 2.4-5.3 mg/kg with minimum concentration in Catfish and maximum concentration in Tilapia. The concentration of nickel in all fish species exceeded WHO permissible limit of 1.40 mg/kg hence consuming fishes from this habitat can cause larynx cancer, asthma and chronic bronchitis. The concentrations of iron and potassium in the fish species were found to be very low. Iron having a range of 1200-1700 mg/kg and potassium having a range of 1100 - 2000 mg/kg. These concentration values are very low hence they pose no significant threat to human health.

The order of concentration of heavy metals in Tilapia and Catfish was given as Ni < Cr < As < Rb < Sr < Co < Se < Ti < Mn < Zn < Cu<K< Fe < Ca respectively, while in carp as Ni <Cr < As < Rb < Sr < Co < Ti < Se< Mn < Zn < Cu< Ca <Fe<K.The concentrations of metallic elements in catfish are lower that of Tilapia and Carp fish., these could be as a result of their feeding habits as Tilipia and Carp fishes are herbivorous while the carp is Carnivorous

The concentrations of some heavy metals in all three fish species were compared with the available FAO/WHO (1989) standard. The results from the tables showed that chromium concentration was quite high when compared to previous study (Uysal et al., 2008). The health risks associated with chromium exposure includes cancer, haemolysis, renal and liver failure and damages to circulatory and nerve tissues. Manganese concentration also exceeded the FAO/WHO (1989) limit of 1.0 mg/kg. Hence, consumers are at risk of having Central nervous system disorder, liver cirrhosis and Parkinson's disease. Cobalt has no FAO/WHO (1989) limit and the health effects associated with high cobalt concentration includes asthma, pneumonia, skin rashes and wheezing (Agency for Toxic Substances and Disease Registry, 2004).The value obtained can be compared to the previous study (Uysal et al., 2008).

The concentration of Nickel fell below the limit set by USFDA (1993) limit of 70-80 mg/kg. Nickel is known to cause larynx cancer, asthma and chronic bronchitis (Sivaperumal et al., 2007). The concentration of copper was also found to exceed the FAO/WHO (1989) limit of 30 mg/kg. Copper is an essential element in the human body but at high concentrations, teratogenicity (Gwozdzinski, 1995) and chromosomal aberrations (Bhunya & Pati, 1987).

Zinc also exceeded the FAO/WHO (1989) limit of 100 mg/kg and is likely to cause zinc toxicosis which in turn causes diarrhoea, bloody urine, liver and kidney failure and anaemia in consumers (Duruibe et al., 2007). Arsenic concentrations in all samples were also higher than the FAO/WHO (1989) limit of 1.4 mg/kg. Large oral doses of Arsenic can result in death. Lower levels of inorganic arsenic cause irritation stomach and intestines, with symptoms such as stomach ache, nausea, vomiting, and diarrhoea. It can also cause abnormal heart rhythm, blood vessel damage and impaired nerve function (Agency for Toxic Substances and Disease Registry, 2007).

Table 2 shows the mean comparison of the level of toxic metals in Gills and Muscle of Tilapia fish. The mean concentration of K in Gills was 1330±200; which was higher than the mean in muscle 1060±0.00. The difference was not statistically significant ($p>0.05$). The difference between the mean Ca in Gills (2294.36±179.66) and the mean in Table 2

Muscle (1377.00±2.65) was statistically significantly ($p<0.05$) higher in the gills than the muscles of Tilapia fish. The amount of Cr in the Gills (5.93±.45) was significantly ($p<0.05$) higher than the muscles (4.27±0.15). The level of Ni in the Gills (4.53±0.67) was higher than that of the muscle (3.57±0.31), the difference in mean was not statistically significant ($p>0.05$). The amount of Fe in the gills of Tilapia fish (1260±50) was significantly lower ($p<0.05$) than the level of Fe in the muscle (1650±50). The Concentration of Zn in the Gills (316.40±4.80) was significantly ($p<0.05$) lower than the concentration in the muscle (360.43±0.25). Se level in the gills (17.76±1.53) was significantly lower than the level of Se in the muscle (23.50±0.30).

Table 3 shows the mean comparison of level of toxic metals in Gills and Muscle of Catfish. The mean concentration of K in Gills was 1.32±0.00; which was lower than the mean in muscle 1.43±0.02. The difference is statistically significant ($p<0.05$). The difference between the mean Ca in Gills (2617.00±4.36) and the mean in Muscle (2158.00±6.00) was statistically significantly ($p<0.05$) higher in the gills than the muscles of Tilapia fish. The amount of Cr in the Gills (4.27±.0.15) was significantly ($p<0.05$) higher than the muscles (3.47±0.35). The level of Ni in the Gills (3.57±0.25) was higher than that of the muscle (3.47±0.35), the difference in mean was statistically significant ($p<0.05$). The amount of Fe

Table 2
Mean comparison for Tilapia fish and with previous study

Parameter	Gills (n=4)	Muscles(n=4)	t	p	Uyal et al., 2009	
					Gills	Muscles
K	1330±200	1060±0.00	24018	0,0730		
Ca	2294.36±179.66	1377.00±2.65	8.843	0.001		
Ti	23.60±2.05	17.70±0.20	4.957	0.008		
Cr	5.93±0.45	4.27±0.15	6.063	0.004	0.39	<DL
Mn	230.91±12.06	204.30±0.56	3.817	0.019	20.70-20700	0.48-480
Fe	1260±50	1650±50	-9.131	0.001	130.60-1718.4	18.44-242.6
Co	12.27±0.97	8.54±0.03	6.643	0.003	0.54-108	<DL
Ni	4.53±0.67	3.57±0.31	2.286	0.084	1.51-503.3	1.51-503.3
Cu	487.73±4.02	397.27±0.50	38.673	0.000	16.07	Yildirim et al. 2009
Zn	316.40±4.80	360.43±0.25	-15.857	0.000	166.75-7579.5	30.06-1366.3
As	7.23±0.15	4.40±0.20	19.500	0.000		
Se	17.67±1.53	23.50±0.30	-6.490	0.003		
Sr	8.87±0.31	4.27±0.15	23.326	0.000		
Rb	7.83±0.35	5.23±0.12	12.182	0.000		

in the gills of Catfish (1.47±0.01) was significantly lower ($p<0.05$) than the level of Fe in the muscle (1.72±0.01). The Concentration of Zn in the Gills (340.77±0.85) was significantly ($p<0.05$) higher than the concentration in the muscle (369.67±2.08). Se level in the gills (16.43±0.15) was significantly higher than the level of Se in the muscle (11.70±0.20). There was no significant difference ($p>0.05$) in the level of Rb in Gills (6.53±0.15) and the muscle (6.67±0.25), although the amount of Rb in the Muscle was higher than the mean concentration in the Gills.

Tables 3 and 4 indicate that on the average, heavy metal concentration in the gills was higher than in muscles for all three species but for few exceptions. This result agreed with the results of previous experiments by (Akpanyung et al., 2014; Etesin & Benson, 2007). The mean concentration of K, Ca, Ti, Cr, Mn, Co, Ni, Cu, As, Sr, Rb in the gills of Tilapia were all higher than the mean concentration in the muscles while the few exceptions were Fe, Zn and Se which were higher in the muscle than in the gills.

Also, the mean concentration of Fe, Ca, Ti, Cr, Mn, Co, Ni, Cu, As, Sr, Rb in the gills of Catfish were all higher than the mean concentration in the muscles while the few exceptions were K, Zn and Se which were higher in the muscle than in the gills. This shows that Zn and Se accumulate more in the muscle than in the gills of both fish species.

The above Table 5 shows the inter-relationship between toxic metals in Tilapia fish. Ca, Ti, Cr, Mn, Co, Sr and Rb had significantly positive relationship with K; while Se had significant negative relationship with K. Ti, Cr, Mn, Co, Cu, As, Sr and Rb had significant positive relationship with Ca; while Fe, Zn and Se were significantly negatively related with Ca. Cr, Mn, Co, Cu, As, Sr and Rb had significantly positive relationship with Ti,

Table 3
Mean comparison for Catfish fish

Parameter	Gills (n=4)	Muscles(n=4)	t	p
K	1320±0.00	1403±20	9.504	0.001
Ca	2617.00±4.36	2158.00±6.00	107.199	0.000
Ti	20.17±0.50	14.53±0.32	16.338	0.000
Cr	4.27±0.15	3.47±0.35	3.618	0.022
Mn	196.67±2.52	166.80±0.46	20.223	0.000
Fe	1470±10	1720±10	34.688	0.000
Co	9.33±0.25	6.67±0.15	15.689	0.000
Ni	3.57±0.25	2.39±0.03	8.035	0.001
Cu	451.67±1.15	369.7±2.08	59.664	0.000
Zn	340.77±0.85	299.90±1.25	46.843	0.000
As	5.27±0.12	3.37±0.31	10.076	0.001
Se	16.43±0.15	11.70±0.20	32.577	0.000
Sr	8.53±0.49	4.63±0.21	12.616	0.000
Rb	6.53±0.15	6.67±0.25	0.784	0.477

Table 4
Mean comparison for Carp fish

Parameter	Gills (n=4)	Muscles(n=4)	T	p
K	26300±0.42	20200±0.00	2.498	0.067
Ca	1339.33±16.17	1117.00±2.00	23.641	0.000
Ti	23.63±1.39	14.73±0.15	11.029	0.000
Cr	5.22±0.01	4.67±0.25	3.784	0.019
Mn	222.33±1.71	215.57±0.25	6.782	0.002
Fe	11300±200	12400±100	-10.165	0.001
Co	13.53±0.34	12.50±0.40	3.410	0.027
Ni	4.20±0.05	3.89±0.06	6.979	0.002
Cu	503.90±1.57	491.77±0.51	12.717	0.000
Zn	309.57±3.65	317.90±0.50	-3.918	0.017
As	7.23±0.01	6.63±0.21	4.985	0.008
Se	19.65±0.17	20.67±0.25	-5.813	0.004
Sr	11.34±0.89	8.27±0.21	-5.824	0.004
Rb	10.05±0.35	8.33±0.25	6.897	0.002

Fe, Zn, and Se had significantly negative relationship with Ti. Mn, Co, Cu, As, Sr and Rb had significantly positive relationship with Cr; while Fe, Zn, Se had significant negative relationship with Cr. Co, Cu, As, Sr and Rb had significant positive relationship with Mn, while Fe, Zn and Se had significant negative relationship with Mn. Zn, Se had significantly positive relationship with Fe, while Co, Ni, Cu, As, Sr and Rb had significant negative relationship with Fe. Ni, Cu, As, Sr and Rb had significantly positive relationship with Co; while Zn and Se had significant negative relationship with Co. Cu, As, Sr and Rb had significant positive relationship with Ni, while Zn and Se had negative significant relationship with Ni. As, Sr and Rb had significant positive relationship with Cu; while Zn and Se had significant negative relationship with Cu. Se had significant positive relationship with Zn; while As, Sr and Rb had significant negative relationship with Zn. Sr and Rb had significant positive relationship with As; while Se had negative significant relationship with As. Sr and Rb had significant negative relationship with Se. Rb had significantly positive relationship with Sr.

Table 6 shows the inter-relationship between toxic metals in Catfish. Ca, Ti, Cr, Mn, Co, Ni, Cu, As, Se and Sr had significantly negative relationship with K; while the Fe had significant positive relationship with K, except Rb which was not significantly related with K. Ti, Cr, Mn, Co, Ni, Cu, Zn, As and Sr had significant positive relationship with Ca; while Fe was significantly negatively related with Ca. Cr, Mn, Co, Ni, Cu, Zn, As and Sr had significantly positive relationship with Ti, Fe was significantly negative relationship with Ti. Mn, Co, Ni, Cu, Zn, As, Se and Sr had significantly positive relationship with Cr; while Fe had significant negative relationship with Cr. Co, Ni, Cu, Zn, As, Se and Sr had

Table 5
Inter-correlation between toxic metals in Tilapia fish

	K	Ca	Ti	Cr	Mn	Fe	Co	Ni	Cu	Zn	**As**	**Se**	**Sr**	**Rb**
K	1	.890*	.951**	.921**	.971**	-.734	.915*	.587	.802	-.726	.778	-.919**	.816*	.859*
Ca	.890*	1	.986**	.991**	.958**	-.940**	.997**	.761	.985**	-.950**	.977**	-.993**	.988**	.996**
Ti	.951**	.986**	1	.989**	.985**	-.885*	.994**	.717	.945**	-.894*	.932**	-.993**	.953**	.974**
Cr	.921**	.991**	.989**	1	.976**	-.930**	.991**	.742	.963**	-.924**	.955**	-.991**	.965**	.979**
Mn	.971**	.958**	.985**	.976**	1	-.870*	.967**	.617	.907*	-.862*	.882*	-.979**	.917**	.943**
Fe	-.734	-.940**	-.885*	-.930**	-.870*	1	-.915*	-.680	-.972**	.987**	-.962**	.924**	-.968**	-.953**
Co	.915*	.997**	.994**	.991**	.967**	-.915*	1	.767	.971**	-.925**	.962**	-.992**	.975**	.988**
Ni	.587	.761	.717	.742	.617	-.680	.767	1	.755	-.679	.792	-.689	.734	.729
Cu	.802	.985**	.945**	.963**	.907*	-.972**	.971**	.755	1	-.987**	.994**	-.969**	.999**	.993**
Zn	-.726	-.950**	-.894*	-.924**	-.862*	.987**	-.925**	-.679	-.987**	1	-.979**	.935**	-.985**	-.970**
As	.778	.977**	.932**	.955**	.882*	-.962**	.962**	.792	.994**	-.979**	1	-.956**	.991**	.984**
Se	-.919**	-.993**	-.993**	-.991**	-.979**	.924**	-.992**	-.689	-.969**	.935**	-.956**	1	-.976**	-.990**
Sr	.816*	.988**	.953**	.965**	.917**	-.968**	.975**	.734	.999**	-.985**	.991**	-.976**	1	.997**
Rb	.859*	.996**	.974**	.979**	.943**	-.953**	.988**	.729	.993**	-.970**	.984**	-.990**	.997**	1

Table 6
Inter-correlation between toxic metals in Catfish

	K	Ca	Ti	Cr	Mn	Fe	Co	Ni	Cu	Zn	As	Se	Sr	Rb
K	1	-.977**	-.967**	-.769	-.978**	.987**	-.984**	-.943**	-.976**	-.971**	-.946**	-.987**	-.976**	.186
Ca	-.977**	1	.994**	.878*	.995**	-.998**	.992**	.970**	1.000**	.999**	.984**	.998**	.988**	-.375
Ti	-.967**	.994**	1	.856*	.979**	-.987**	.979**	.986**	.995**	.989**	.992**	.987**	.972**	-.386
Cr	-.769	.878*	.856*	1	.885*	-.857*	.849*	.811	.876*	.894*	.879*	.859*	.866*	-.619
Mn	-.978**	.995**	.979**	.885*	1	-.996**	.993**	.943**	.993**	.996**	.972**	.997**	.994**	-.333
Fe	.987**	-.998**	-.987**	-.857*	-.996**	1	-.996**	-.962**	-.997**	-.997**	-.972**	-1.000**	-.991**	.333
Co	-.984**	.992**	.979**	.849*	.993**	-.996**	1	.944**	.991**	.990**	.970**	.996**	.998**	-.340
Ni	-.943**	.970**	.986**	.811	.943**	-.962**	.944**	1	.973**	.965**	.963**	.959**	.927**	-.393
Cu	-.976**	1.000**	.995**	.876*	.993**	-.997**	.991**	.973**	1	.998**	.987**	.997**	.987**	-.374
Zn	-.971**	.999**	.989**	.894*	.996**	-.997**	.990**	.965**	.998**	1	.979**	.996**	.987**	-.385
As	-.946**	.984**	.992**	.879*	.972**	-.972**	.970**	.963**	.987**	.979**	1	.974**	.969**	-.415
Se	-.987**	.998**	.987**	.859*	.997**	-1.000**	.996**	.959**	.997**	.996**	.974**	1	.993**	-.323
Sr	-.976**	.988**	.972**	.866*	.994**	-.991**	.998**	.927**	.987**	.987**	.969**	.993**	1	-.342
Rb	.186	-.375	-.386	-.619	-.333	.333	-.340	-.393	-.374	-.385	-.415	-.323	-.342	1

significant positive relationship with Mn, while Fe had significant negative relationship with Mn. Co, Ni, Cu, Zn, As, Se and Sr had significantly negative relationship with Fe. Ni,Cu, Zn, As, Se and Sr had significantly positive relationship with Co. Cu, Zn, As, Se and Sr had significant positive relationship with Ni. Zn, As, Se and Sr had significant positive relationship with Cu. As, Se and Sr had significant positive relationship with Zn. Se and Sr had significant positive relationship with As. Sr had significant negative relationship with Se. The correlation value showed in Table 5 and 6 could be that the fishes were caught from the same source

The above Table 7 shows the inter-relationship between toxic metals in Catfish. Majority of the relationships between toxic metals in Carp were not statistically significant. Zn showed a significant positive relationship with Ca. Mn showed a significant negative relationship with Ti. Co showed a positive significant relationship with Fe. Cu showed a negative significant relationship with As. Se showed a significant negative relationship with Rb.

Figure 2 is a box plot showing the distribution of pH and EC of the water.in this study. The median pH recorded in this study is 10.34, with an Interquartile Range (IQR) of 10.2-11.16; while the median EC is 2.77×10^3, with an IQ Range of $2.39\text{-}3.16 \times 10^3$. The distribution did not show any outliers in both the pH and EC.

Figure 3 shows the hazard index of mixtures of metal intake from consumption of fishes from Oluwa river for adults and children. In children the HI value was highest in Carp (43.1451) and least in Catfish (32.8077) while for adults HI value was highest in Carp (14.9726) and least in Catfish (11.2484) as seen in Figure 3.

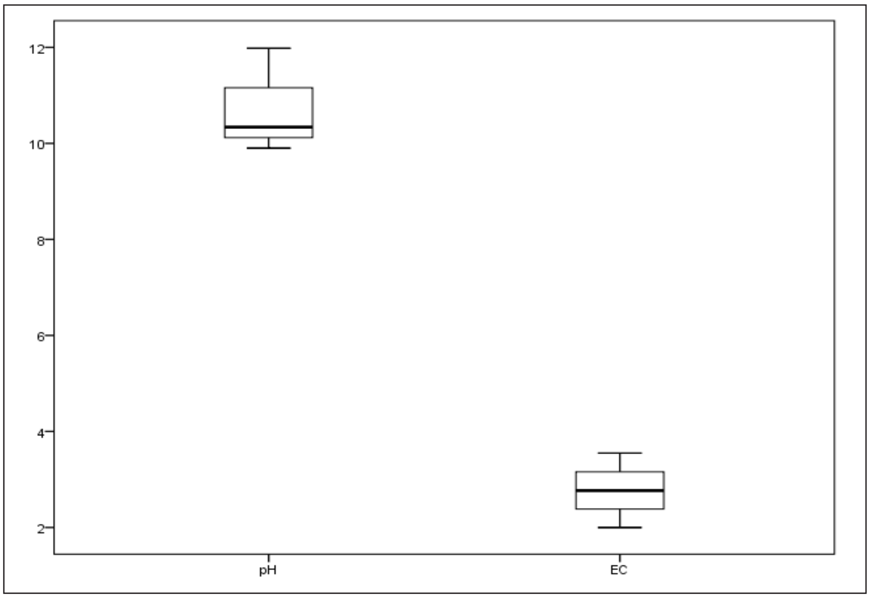


Figure 2. Box plot showing the physico-chemical characteristics of the water

Table 7
Inter-correlation between toxic metals in Carp

	K	Ca	Ti	Cr	Mn	Fe	Co	Ni	Cu	Zn	As	Se	Sr	Rb
K	1	-.982	.929	.737	-.954	.277	.327	-.556	-.842	-.982	.839	-.434	.891	.434
Ca	-.982	1	-.982	-.596	.993	-.454	-.500	.703	.724	1.000**	-.721	.596	-.961	-.596
Ti	.929	-.982	1	.434	-.997*	.614	.655	-.825	-.581	-.982	.577	-.737	.996	.737
Cr	.737	-.596	.434	1	-.500	-.445	-.397	.151	-.985	-.596	.986	.289	.350	-.289
Mn	-.954	.993	-.997*	-.500	1	-.553	-.596	.780	.640	.993	-.636	.684	-.986	-.684
Fe	.277	-.454	.614	-.445	-.553	1	.999*	-.953	.286	-.454	-.291	-.986	.683	.986
Co	.327	-.500	.655	-.397	-.596	.999*	1	-.967	.235	-.500	-.240	-.993	.721	.993
Ni	-.556	.703	-.825	.151	.780	-.953	-.967	1	.020	.703	-.014	.990	-.873	-.990
Cu	-.842	.724	-.581	-.985	.640	.286	.235	.020	1	.724	-1.000**	-.122	-.505	.122
Zn	-.982	1.000**	-.982	-.596	.993	-.454	-.500	.703	.724	1	-.721	.596	-.961	-.596
As	.839	-.721	.577	.986	-.636	-.291	-.240	-.014	-1.000**	-.721	1	.127	.500	-.127
Se	-.434	.596	-.737	.289	.684	-.986	-.993	.990	-.122	.596	.127	1	-.795	-1.000**
Sr	.891	-.961	.996	.350	-.986	.683	.721	-.873	-.505	-.961	.500	-.795	1	.795
Rb	.434	-.596	.737	-.289	-.684	.986	.993	-.990	.122	-.596	-.127	-1.000**	.795	1

The Target Hazard Quotient (THQ) of each metal through consumption of fishes from Oluwa river for both adults and children increased in the following order: Fe < Ni < Cr < Cu <Mn< As. The maximum value of THQ was seen in carp (17.96 and 6.16) for adults and children in As (Figures 4 and 5). The THQ values of other fish species varied from Cu (2.86-3.42), Cr (0.72-0.95), As (4.01-6.16), Ni (0.041-0.056), Fe (0.0005–0.006), and Mn (3.61−4.33), respectively, for adults and also for children ranged from Cu (8.43-9.9), Cr (2.09-2.76), As (11.7-17.9), Ni (0.12-0.16), Fe (0.0014-0.0017), and Mn (10.5-12.6), respectively. The THQ of Cu, As and Mn were all less than 1 in all species for adults while for children, all metals except Ni and Fe had their THQ greater than 1 for all the fish species.

The THQ values for Cr, Ni and Fe for adults and that of Ni and Fe for children were below one in all three fish species. Hence, the intake of these metals by consuming these fishes is not likely to cause any appreciable health risk. While in the case of Cu, As and Mn in adults and Cu, As, Mn and Cr in children, the THQ for all three fish species exceeded the limit of one as shown in Figure 5. This indicates that there is potential health risks associated with the intake of these metals through consumption of these fishes. Copat et

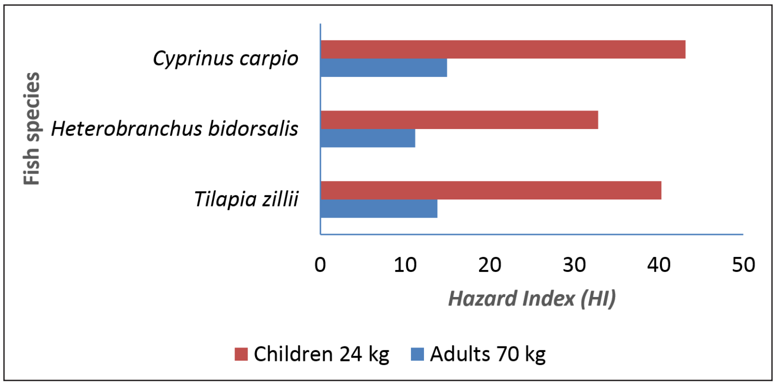


Figure 4. Target hazard quotient (THQ) in children exposed to fishes from Oluwa River

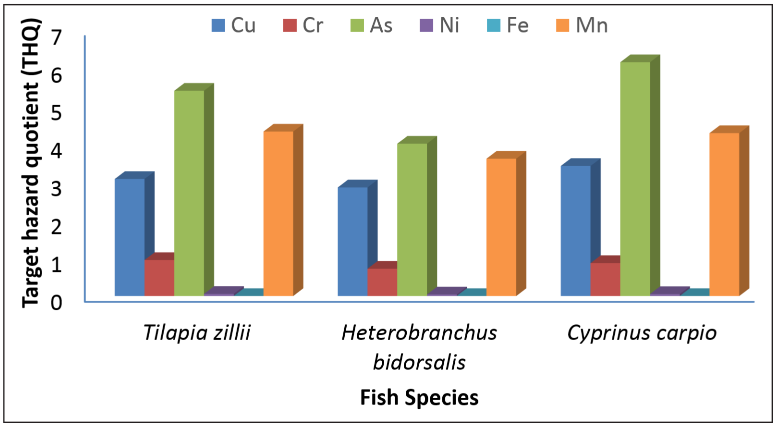


Figure 5. Target hazard quotient (THQ) adults exposed to fishes from Oluwa River

al. (2013) and Yabanli and Alparslan (2015) reported that the THQ values for Cr, Mn, Ni, V, Cu and Zn were below 1 .The values of THQ in Cr and Cu were in agreement with the values reported by previous studies (Copat et al., 2013; Yabanli & Alparslan, 2015). The THQ value has been recognized as one of the reasonable parameters for the risk assessment of metals associated with the consumption of contaminated fish (Li et al., 2013).

CONCLUSION

The concentrations of Zn, Fe, Mn and Cu in the fish samples from Oluwa River were higher than the FAO/WHO limit. The concentrations of metallic elements in Tilapia and Carp fishes were higher than the elemental concentrations in Cart fish except in Ca, Ti and Fe. These differences in metallic concentrations could be as a result of feeding habits of the fishes. Tilapia and Carp fishes are herbivorous in feeding while cart fish is carnivorous. The THQ of Cu, As and Mn were all less than 1 in all species for adults while for children, all metals except Ni and Fe had their THQ greater than 1 for all the fish species.The HI value in children was highest in Carp and least in Catfish while for adults HI value was highest in Carp and least in Catfish. The result of this study could also establish a baseline for future studies of heavy metal pollution.

RECOMMENDATION

There is a need for further extensive study and particularly the accumulation of heavy metals in humans.Point sources of heavy metals in the waters especially run-offs from small scale industries, farmland and indiscriminate dumping of waste in the water body should be closely monitored & proper treatment before disposal into water bodies should be enforced. Continuous monitoring of this water body should be done in order to ensure that the measures put in place to reduce the heavy metals concentrations in the fishes in Oluwa River, Okitipupa is reduced. Level of Heavy metal concentration in water bodies in the neighboring environments of Okitipupa should also be accessed for health and safety reasons.

ACKNOWLEDGEMENT

We appreciate the Dean Faculty of Science and the Head of Department Chemical Science for their logistic support and prayers which have really been vital in the success of this work. We also appreciate the fishermen for their immeasurable support.

REFERENCES

Aderinola, O. J., Clarke, E. O., Olarinmoye, O. M., Kusemiju, V., & Amatekhai, M. A. (2009). Heavy metals in surface water, sediments, fish and Periwinkles of lagos lagoon*. American-Eurasian Journal of Agricultural and Environmental Sciences, 5*(5), 609-617.

Agency for Toxic Substances and Disease Registry. (2004). *Toxicological Profile for Cobalt*. Atlanta, GA: U.S. Department of Health and Human Services, Public Health Service.

Agency for Toxic Substances and Disease Registry. (2007). Toxicological Profile for Arsenic. Atlanta, GA: U.S. Department of Health and Human Services, Public Health Service.

Akpanyung, E. O., Akanemesang, U., Akpakpan, E. I., & Anodoze, N. (2014). Levels of heavy metals in fish obtained from two fishing sites in Akwa Ibom State, Nigeria. *African Journal of Environmental Science and Technology, 8*(7), 416-421.

Awadesh, N. J. (2004). Genotoxicological studies in aquatic organisms: an overview. *Mutation Research, 55*(2), 1-7.

Bhunya, S. P., & Pati, C. P (1987) Genotoxicity of an inorganic pesticide, copper sulphate, in a mouse in vivo test system. *Cytologia, 52*, 801-808.

Copat, C., Arena, G., Fiore, M., Ledda, C., Fallico, R., Sciacca, S., & Ferrante, M. (2013). Heavy metals concentrations in fish and shellfish from Eastern Mediterranean Sea: consumption advisories. *Food and Chemical Toxicology, 53*, 33-37.

Dawson, E. J., & Macklin, M. G. (1998). Speciation of heavy metals in floodplain and flood sediments: a reconnaissance survey of the Aire Valley, West Yorkshire, Great Britain. *Environmental Geochemistry and Health, 20*(2), 67-76.

Duffus, J. H. (2002). "Heavy metals" a meaningless term? (IUPAC Technical Report). *Pure and applied chemistry, 74*(5), 793-807.

Duruibe, J. O., Ogwuegbu, M. C., & Egwurugwu, J. N. (2007). Heavy metal pollution and human biotoxic effects. *International Journal of Physical Sciences, 2*, 112-118.

Dybem, B. (1993). Field sampling and preparation subsamples of aquatic organism for analysis metals and organochlorides, FAO. *Fishery Technology, 212*, 1-13.

Edward, J. B., Idowu, E. O., Oso, J. A., & Ibidapo, O. R. (2013). Determination of heavy metal concentration in fish samples, sediment and water from Odo-Ayo River in Ado-Ekiti, Ekiti-State, Nigeria. *International Journal of Environmental Monitoring and Analysis, 1*(1), 27-33.

Etesin, M. U., & Benson, N. U. (2007). Cadmium, copper, lead and zinc tissue levels in Bonga Shad (Ethmalosa fimbriata) and Tilapia (Tilapia guineensis) caught from Imo River, Nigeria. *American Journal of Food Technology, 2*(1), 48-54.

FAO/WHO. (1989). *National Research Council Recommended Dietary Allowances* (10th Ed.). Washington, DC: National Academy Press.

FAO/WHO. (2010). *Codex Alimentarius Commission Procedural Manual* (19th Ed.). Rome, Italy: World Health Organization and Food and Agriculture Organization of the United Nations.

Fergusson, J. E. (1990). *The Heavy Elements: Chemistry, Environmental Impact and Health Effects*. Oxford, England: Pergamon Press.

Gwozdzinski, K. (1995). Structural changes in erythrocyte components induced by copper and mercury. *Radiation Physics and Chemistry, 45*(6), 877-882.

Jarup, L. (2003). Hazards of heavy metal contamination. *British Medical Bulletin, 68*(1), 167-182.

Kakulu, S. E., Osibanjo, O., & Ajayi, S. O. (1987). Trace metal content of fish and shell fishes of the Niger-Delta area of Nigeria. *Environment International, 13*(3), 247-251.

Li, J., Huang, Z., Hu, Y., & Yang, H. (2013). Potential risk assessment of heavy metals by consuming shellfish collected from Xiamen, China. *Environmental Science and Pollution Research, 20*(5), 2937-2947.

Maureen, A. (2013). Nigeria under pressure to meet 700,000 metric tonnes fish deficit. *Punch Newspaper, Nigeria, 2*, 3-15.

Narain, N., & Nunes, M. L. (2007). *Handbook of meat, poultry and seafood quality*. Iowa, USA: Blackwell Publishing.

Odeemelam, S. A., Iwuozor, C. C., & Ozuo, J. U. (1999, May 18-20). Baseline levels of some toxic metals in selected Nigerian rivers. In *2nd National Conference of the tropical Environment Forum (TEFO)*. Calabar, Nigeria.

Oti-Wilberforce, J. O., Nworie, F., & Oti-wilberforce, R. O. (2016). Comparative assay of heavy metals in gills and muscles of common fishes in Afikpo, Ebonyi State, Nigeria. *World Applied Sciences Journal, 34*(2), 268-272.

Singh, A., Sharma, R. K., Agrawal, M., & Marshall, F. M. (2010). Risk assessment of heavy metal toxicity through contaminated vegetables from waste water irrigated area of Varanasi, India. *Tropical Ecology, 51*(2), 375-387.

Sivaperumal, P., Sankar, T., & Nair, P. V. (2007). Heavy metal concentrations in fish, shellfish and fish product from internal markets of India vis-a-vis international standards. *Food Chemistry, 102*(3), 612-620.

Szentmihalyi, K., & Then, M. (2007). Examination of microelements in medicinal plants of the Carpathian basin. *Acta Alimentaria, 36*, 231-236.

Terra, B. F., Araujo, F. G., Calza, C. F., Lopes, R. T., & Teixeira, T. P. (2008). Heavy metal in tissue of three fish species from different trophic levels in a tropical Brazilian river. *Water Air Soil Pollution, 187*, 275-284.

Trichopoulos, D., (1997). Epidermilogy of cancer. In V. T. DeVita (Ed.), *Cancer, principles and practice of oncology*. Philadelphia: Lippincott Company.

Tüzen, M. (2003). Determination of heavy metals in fish samples of the Mid Dam Lake Black Sea (Turkey) by graphite furnace atomic absorption spectrometry. *Food Chemistry,* 80, 119-123.

USFDA. (1993). *Food and drug administration, guidance document for nickel in shell fish*. Washington, DC: DHHS/PHS/FDA/CFSAN/office of seafood.

USEPA. (1999). *Particulate Matter (PM2.5) Speciation Guidance: US Environmental Protection Agency* (Final Draft). United States Environmental Protection Agency.

USEPA. (2011). *USEPA Regional Screening Level (RSL) Summary Table*. United States Environmental Protection Agency.

USEPA. (2015). *Regional ScreeningLevel (RSL) Summary Table*. United States Environmental Protection Agency.

Uysal, K., Köse, E., Bülbül, M., Dönmez, M., Erdoğan, Y., Koyun, M., ... & Özmal, F. (2008) The comparison of heavy metal accumulation ratios of some fish species in Enne Dame Lake (Kütahya/Turkey). *Environmental Monitoring and Assessment, 157*(1-4), 355-362.

Yabanli, M., & Alparslan, Y. (2015). Potential health hazard assessment interms of some heavy metals determined in demersal fishes caught in eastern Aegean Sea. *Bulletin of Environmental Contamination and Toxicology, 95*(4), 494-498.

Yildirim, Y., Gonulalan, Z., Narin, I., & Soylak, M. (2009). Evaluation of trace heavy metal levels of some fish species sold at retail in Kayseri, Turkey. *Environmental Monitoring and Assessment, 149*(1-4), 223-228.

Pertanika J. Sci. & Technol. 27 (4): 2221 - 2243 (2019)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Fatty Acid Profile and Production of Fatliquor from *Canarium schweinfurthii* Mesocarp Oil

Adachukwu Nwaka Nkwor[1]*, Pius Oziri Ukoha[2], William Robert Wise[3], Njemuwa Njoku Nwaji[1] and Karl Flowers[3]

[1]*Department of Chemistry/Biochemistry/Molecular Biology, Alex Ekwueme Federal University Ndufu-Alike Ikwo, P.M.B 1010, Abakaliki, Ebonyi State, Nigeria*
[2]*Department of Pure and Industrial Chemistry, University of Nigeria, Nsukka 410001, Enugu State, Nigeria*
[3]*Institute for Creative Leather Technologies, The University of Northampton, Park Campus, Boughton Green Road, NN2 7AL, Northampton, United Kingdom*

ABSTRACT

Fatty acid profile of *Canarium schweinfurthii* mesocarp oil was determined by GC-MS. Sulphonated fatliquor was synthesized from the oil and characterized by FT-IR, ^{1}H NMR, ^{13}C NMR, and DSC. The fatliquor was applied onto light leather in the processing of leather shoe upper and physical tests carried out on the fixed leather. The sulphonated *C. schweinfurthii* mesocarp oil had good characteristics as a leather fatliquor as shown by the physical and strength properties of the fatliquored leather. In addition, a significantly opened up structure of the leather treated with the prepared sulphonated oil was observed as indicated from the Scanning Electron Microscopy (SEM) images. The features of the processed trial leathers were comparable with similar leather made with commercially available fatliquor.

Keywords: Canarium schweinfurthii, characterisation, fatliquor, leather, oil, sulphonated

ARTICLE INFO

Article history:
Received: 31 January 2019
Accepted: 11 July 2019
Published: 21 October 2019

E-mail addresses:
adannkwor@yahoo.com; adachukwu.nkwor@funai.edu.ng (Adachukwu Nwaka Nkwor)
pius.ukoha@unn.edu.ng (Pius Oziri Ukoha)
will.wise@northampton.ac.uk (William Robert Wise)
njemuwa2005@yahoo.com (Njemuwa Njoku Nwaji)
plugsly@gmail.com (Karl Flowers)
* Corresponding author

INTRODUCTION

In the leather industry, many chemical and mechanical processes are needed to produce finished leather. Fatliquoring process, the last of the wet processing stage in leather manufacturing is a significant step and its importance cannot be overemphasized. When water is removed from chrome tanned leather during the drying stage, cohesion of the fibres takes place and results in leather

ISSN: 0128-7680
e-ISSN: 2231-8526

which is dry, hard, intractable and difficult to work with (Burgess, 1993). The introduction of fatty matter as lubricant in finely dispersed form in a water medium carried out during the fatliquoring process keeps the fibres apart during drying and reduces the frictional forces within the fibre weave, thereby allowing the fibres to slide over one another. Such lubricants include: natural lubricants which are made up of a blend of natural plant / animal oils and emulsifying agents; sulphonated lubricants obtained from a reaction of sulphuric acid and a double bond in an unsaturated fatty acid; sulphited lubricants, produced when highly oxidised fatty raw materials react with sodium bisulphite to give oxy-sulphonic acids. Other lubricants include synthetic lubricants, produced when various kinds of chemical modifications such as sulpho-chlorination are catalytically carried out on paraffins (Daniels, 2001).

These lubricants, known as fatliquors, also improve the mechanical/strength properties of leather such as tensile strength, pliability, soft handle and tear load (Heidemann, 1993). They also aid resistance to acid attack and help in preventing the ugly appearance of chrome tanned leather after drying (Janakiram, 2007).

The property obtained in each vegetable-based fatliquor differs from the other and it depends on the properties found in the vegetable oil used (Janakiram, 2007). *Canarium schweinfurthii*, also known as black date, is an exotic tree which is abundantly available in Sub-Sahara Africa and can be readily found in African countries such as Nigeria, Ethiopia, Sierra-Leone and Sudan (Evans, 2004). The tree, which grows in the wild in Africa, produces oval- shaped fruits. These fruits have edible mesocarps which are eaten as supplement to diet.

In this study, oil from the mesocarp of *Canarium schweinfurthii* was sulphonated and used as a leather lubricant for chrome tanned goatskins. After the characterisation of the oil and the sulphonated fatliquor had been produced, the fatliquoring ability was assessed by comparing it with an imported fatliquor commonly used in Nigerian tanneries.

MATERIALS AND METHODS

Sample Preparation of Fruits of *C. schweinfurthii*

Mature fruit samples of *C. schweinfurthii* were obtained from where they grow wildly at Onueke, in Ebonyi State, Nigeria. The samples were identified in the Herbarium Unit, Department of Biological Sciences, Ahmadu Bello University, Zaria, Kaduna State, Nigeria and voucher number 7232 was assigned to it. The mesocarp was manually peeled from the fruits to separate it from the kernel and dried in the oven (40^oC) for five days until it became crispy, thus obtaining minimum moisture content (Abayeh et al., 1999). The dried mesocarps were coarsely ground (approximately 2 mm) using the Corona hand kitchen grinder, before extraction. The extraction of the oil from the seeds was carried out in a soxhlet apparatus using n-hexane as a solvent.

Physical and Chemical Characteristics

Both the sulphonated and unsulphonated oils were analyzed for their physical and chemical properties according to the American Oil Chemists Society Methods (Firestone, 1998).

Fatty Acid Composition

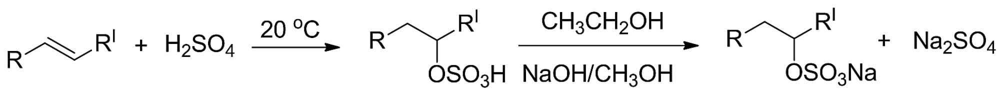
Fatty-acid composition of *Canarium schweinfurthii* oil, CSO was determined using its methyl ester prepared with the method described by Adewuyi et al., (2014) on an Agilent19091S-433HP-5MS gas chromatograph attached to a mass spectrometer. The injection and detection temperatures were 280 and 300°C respectively. Helium was used as the carrier gas at a flow rate of 20 ml/min. The area percentages were recorded with a standard Chemstation Data system. For the mass spectrometry, an ACQ mode scanner (with scan range of 15-500 amu and voltage of 2094) was used and the mass spectra were compared with the NIST11 mass spectral library.

Sulphonation Process

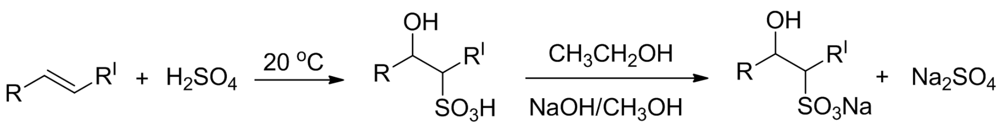
Concentrated sulphuric acid (45 ml) was added dropwise into 150 g of *C. schweinfurthii* oil (with constant stirring at 20°C for 2 h). The crude mass was dissolved in 450 ml of ethanol, and neutralised using 15% NaOH (solubilised in methanol). The salts were filtered off under vacuum. The solvent was removed and recovered using a rotary evaporator (Steik, 1943). The resulting sulphonated product was ready for use as a leather fatliquor.

This reaction of unsaturated oils with sulphuric acid is additional.

Sulphonation process:



Scheme 1. Sulphonation of *C. schweinfurthii* oil to produce sulphonated (Sulphated) oil



Scheme 2. Side by side reaction of the sulphonation of *C. schweinfurthii* oil

Melting Point Determination

The thermal behaviour of the *C. schweinfurthii* oil, CSO and Sulphonated *C. schweinfurthii* oil, SCSO was determined using the Mettler DSC 2 Star System in temperature range of -80 to 180ºC.

FT-IR Analysis

The changes in the functional groups of the oils were studied using the FT-IR analysis by FT-IR measurement (600-4000 cm^{-1}), normal resolution of 4 cm^{-1} using a Shimadzu 8400S FT-IR instrument (Shimadzu, Milton Keynes, UK).

NMR Analysis

^1H nuclear magnetic resonance (NMR) and ^{13}C NMR spectra of both CSO and SCSO were acquired on a Bruker Biospin AV500 – 5mm BBO probe with Z axis gradient, TOPSPIN v 2.1, ^1H=500.13 MHz, ^{13}C=125.76 MHz (Brucker, Coventry, UK).

Fatliquoring Process

Wet blue goat skin, shaved at 1.2-1.3 mm, was divided into four quarters such that the sampling positions **(**BS 2418, 2002), were uniformly represented in all the four quarters.

Further treatments on each of the four quarters of the wet blue goat skins labelled NC, PC, A1 and A2 respectively, were simultaneously carried out *(with the aid of four separate tanning drums)* using a conventional shoe upper manufacturing process [fatliquoring process] (ICLT SR15/31, 2015).

A negative control (designated NC) was processed without any fatliquor; a positive control (designated PC) was processed using a commercial sulphated fatliquor, Trupon DXV (Trumpler Gmbh, Worms, Germany). Sample A1 was processed using pure sulphonated *C. schweinfurthii* oil; Sample A2 was processed using a blend of pure sulphonated *C. schweinfurthii* oil and 7.5% raw castor oil. Leather dyeing was omitted in the process to enable the Sudan IV staining test (for identification of fatty substances) be carried out effectively after the leather manufacturing. The fatliquoring process used was illustrated in Table 1.

Strength/Mechanical Properties of Leather

These leather samples: NC, PC, A1 and A2 were conditioned according to (BS 2419, 2002) before mechanical tests were carried out on them. The samples were staked twice using a Cartigliano PAL 160 leather staking machine (Cartigliano, Bassano, Italy) on the lowest setting. The Strength / mechanical properties of leather samples were all determined using; softness (BS 1723, 2015) tensile strength (BS 3376, 2011), elongation at break and tear strength of leather (BS 3377, 2011) and grain strength standards (BS 3379, 2015).

Sudan Stain Test on Leather Samples

Thin sections (50 µm) of the leather samples were cut with a Leica 1850 cryostat microtome (Leica, Wetzler, Germany) (set at -20°C) and used in the Sudan (IV) stain test (Bugby, 1999) for the determination of the extent of penetration of the fatliquors into the leather fibrils.

Scanning Electron Microscopy

Cross sections of experimental and control were examined for changes in fibre structure using Hitachi S-3000N scanning electron microscope, SEM (Hitachi, Maidenhead, UK).

Leather Samples were gold-coated using an Emscope SC 500 gold sputter coater before SEM analysis (Quorom Technologies, Laughton, UK).

Table 1
Fatliquoring process for shoe upper manufacture damp shaved weight: 400 g

Process	% (m/m)	Chemical	T(°C)	Time(min)	Comments
Wet Back	300	Water	30		
	0.2	Surfactant		20	
Drain					
Neutralise	100	Water	35		-
Add	1	Sodium formate		5	
	0.25	Sodium bicarbonate		30	pH = 4.6 Bromo-cresol green cross-section = yellow/ green
Drain					
Wash	200	Water	35	5	
Drain					
Retan/Fat	100	Water	30		
Add	6	Replacement syntan		15	
	4	Vegetable tannin		30	pH = 4.63
	20	Water	35		
	3	Acrylic resin		30	pH = 5.29
Drain					
Wash	200	Water	50	5	
Drain					
Dye/Fat	100	Water	50		
Add	2	Dye		10	Paste if necessary
	6	Fatliquor (1:3)		40	Run longer if needed
Fix add	1	Formic acid (1:10)		20	pH =3.6
Drain, wash X2, horse up					

Source: ICLT SR 15/31

RESULTS AND DISCUSSIONS

Physicochemical Properties

The physicochemical properties of CSO and SCSO are shown in Table 2. The high percentage quantity of oil in *C. schweinfurthii* indicates potential applicability for possible large-scale chemical modification in the synthesis of fatliquor via sulphonation.

The green colour possessed by CSO was not a disadvantage to the final colour of the leather article produced as shown by the 10% solution that had a pale green colour. The specific gravity of the oil is in line with the density of most vegetable oils (Gunstone, 2004). The translucent (not so clear) nature of the 10% solution is an indication of a medium degree of sulphonation (Waite, 1999). This physical observation shows that there was a percentage of SO_3 (anionic emulsifier) incorporated into the sulphonated compound. It should also be noted that SO_3 is the fuel which drives the oil droplets into the leather. It equally ensures a great degree of fixation as they will be attached to the positively charged leather (Covington, 2011).

Fatty Acid Composition

The nature of the fatty acids present in oils affects the ease of sulphonation and the nature of sulphonated product obtained. It can be seen from Table 3 that the ratio of the total unsaturated fatty acids of *C. schweinfurthii* is almost equal to the total saturated fatty acids (1: 0.80). Palmitic acid constitutes more than 94% of the saturated fatty acids while linoleic acid constitutes more than 54% of the unsaturated fatty acid. Oleic acid is also a major unsaturated fatty acid constitutes more than 43% of its content. Details of the GC-MS analysis are shown in Appendix A, B, C, D, E, F, G, H. Since a great percentage of unsaturated fatty acids (55.53%) is found in the oil, the double bonds present in them are readily available for the sulphonation reaction.

Table 2
Physicochemical properties of both CSO and SCSO

Parameter	CSO	SCSO
Percentage Yield (%)	51.32	69.7
Colour	Dark green	Dark green
Odour	Inoffensive	Odourless
Appearance of 10% Solution	-	Translucent
Colour of 10% solution	-	Light green
pH of 10% Solution	-	7.44
Stability of 10% solution	-	Stability> 24hrs
%Ash Content	-	Trace
% SO_3	-	4.05
Specific gravity (at 20°C)	0.948	0.978
Acid Value (mg KOH g^{-1})	11.36	8.25
Free fatty acid (as oleic acid)	5.68	4.13
Iodine value	68	23
Saponification value	196	192

Table 3
Fatty acid profile of CSO

Fatty acid	Percentage composition (%)
Palmitic acid	41.81
Stearic acid	2.66
Saturated fatty acids	44.47
Oleic acid	23.75
Linoleic acid	30.08
Palmitoleic acid	1.70
Unsaturated fatty acids	55.53

Differential Scanning Calorimetry (DSC) Results

It was observed from the DSC results that CSO had a wide melting range of (0.64 to 8.89^{o}C) depicting the various fatty acids (saturated and unsaturated) present in the oil (Table 4). These melting ranges and DSC curves result from the combined effects between fatty acid composition, polymorphism of natural oils and fats, and thermal history (Kaiserberger, 1989). The melting behaviour of the oils CSO and SCSO studied varied due to the different characteristics and compositions of fatty acids present in the triglycerides (Fashina et al., 2008**.**

The increased melting point observed in the SCSO shows that most of the unsaturated fatty acids have been used up in the sulphonation reaction; leaving behind saturated fatty acids (which have a higher melting point) than unsaturated fatty acids (Berg et al., 2002).

Emulsion Stability Tests

Table 5 shows that 10% fatliquor emulsions of SCSO are generally stable in various salt solutions used in leather manufacturing processes like deliming, pickling and chrome tanning steps. When sulphonated fatliquor which is anionic ionizes in water to release the anionic sulphonate group, these negative ions react electrostatically with the protonated amino group from the basic group of protein in leather and thus fixed up with leather.

This fixation can only be done in an acid medium. The SCSO has a good emulsion stability towards tanning salts, pickle liquor and hard water, but unstable when in contact with formic acid. This instability of the emulsion with formic acid enables the fatliquor to stay in the leather and form bonds that hold it there (Habib & Alshammari, 2014). Table 5, therefore, shows that SCSO can also be used in the retanning and fatliquoring steps.

Table 4
The thermal behaviour of CSO and SCSO

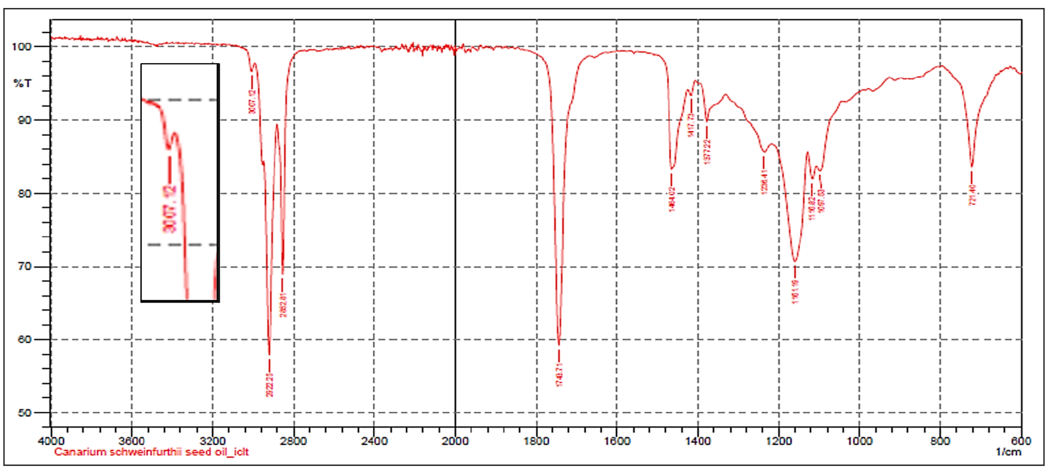
Oil Sample	Onset Temperature (°C)	Peak Temperature (°C)	Endset Temperature (Melting Point) (°C)
CSO	0.64	4.42	8.89
SCSO	7.54	16.5	18.44

Table 5
Stability of 10% fatliquor emulsion of SCSO towards pickle liquor, tan liquor and hard water

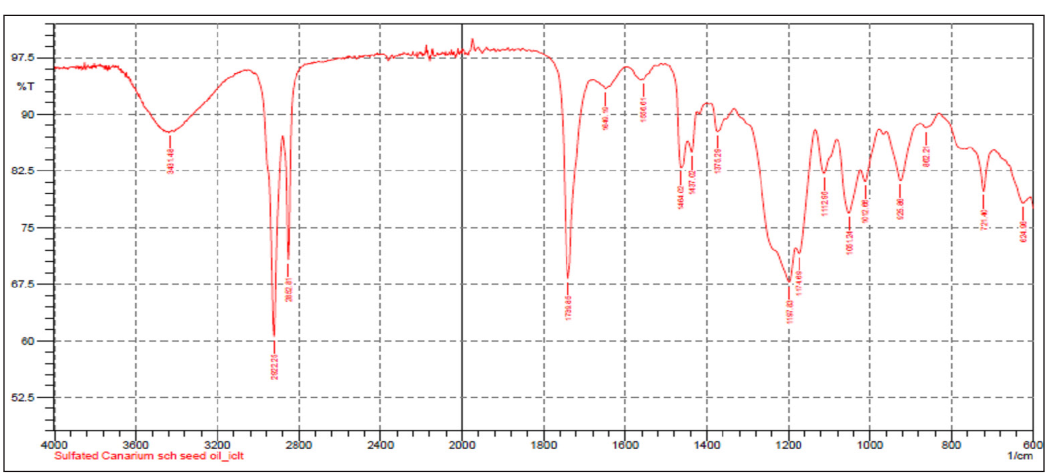
Solution added	Stability Status
5 % Basic chromium sulfate (tan liquor)	Stable
5 % $MgSO_4$ (hard water)	Stable
5 % NaCl (found in pickle liquor)	Stable
5 % Formic acid	Unstable

Fourier Transform Infra-Red (FT-IR) Results

The FT-IR results in Figure 1a and 1b illustrate the IR spectrum of CSO and SCSO. In Figure 1b, the probable attack of H_2SO_4 on the –C=C- to form the sulphonated product was confirmed by the absence of unsaturated C-H stretching peak at 3007.12 cm^{-1} in



(a)



(b)

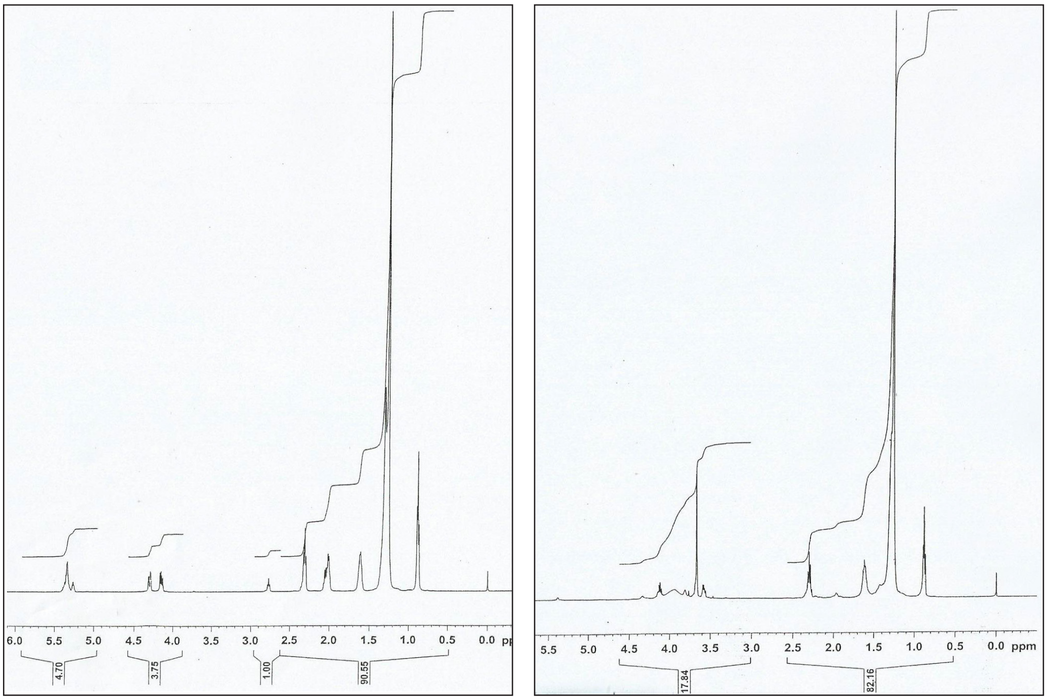
Figure 1. FT-IR spectra of CSO (a) and SCSO (b) (Insert: expanded section of signal showing 3007.12 cm^{-1})

non-conjugated. In addition, the presence of vibrational peak at 1198 cm^{-1} (Figure 1b) corresponds to S=O stretching of sulphonated groups, which is absent in Figure 1a; hence, confirming the formation of sulphonated product.

Other prominent peaks found in both samples are at (2853 cm^{-1}) C-H stretching frequency of alkane; (1744 cm^{-1} (C=O stretching frequency of ester); 1464 cm^{-1} (bending frequency of unsaturated alkene); and 721 cm^{-1} (bending frequency of saturated carbon atom). The OH peak at 3431 cm^{-1} in Figure 1b depicts the traces of alcohol used in the formation of the sulphonated products.

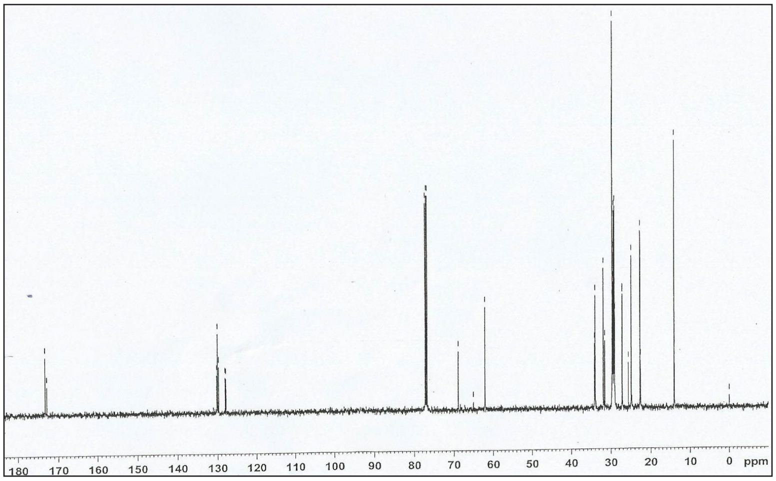
Nuclear Magnetic Resonance (NMR) Spectroscopy Results

The ^1H NMR of CSO and SCSO is shown in Figures 2a and 2b respectively, while the ^{13}C NMR spectral diagrams are found in Figures 3a and 3b respectively. The ^1H NMR displayed the unsulphonated multiplet olefinic protons attached to C=C double bond at δ 5.29 ppm, integrated into 5 protons. This is expected since these protons are sp^2 hybridized, resulting in deshielded NMR signals due to the influence of the diamagnetic anisotropy of the π system. Sulphation / sulphonation usually lead to the saturation of the double bond. The sp^3 hybridized protons formed are thus expected to be shielded relative to the sp^2 olefinic protons. The newly formed protons (H-C-S and H-C-O) in the SCSO showed signals at

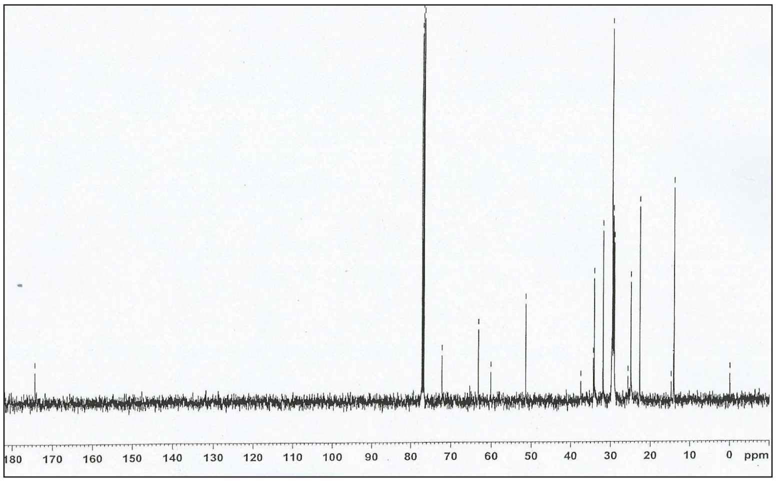


(a) (b)

Figure 2. ^1H NMR of CSO (a) and SCSO (b) in deuterated chloroform



(a)



(b)

Figure 3. ^{13}C NMR of CSO (a) and SCSO (b) in deuterated chloroform

δ 3.6 and 3.73 ppm. It is important to note that the slight deshielding observed for these protons relative to the rest of the protons in the SCSO is due to the inductive effect of the electronegative sulphur and oxygen atoms. The inductive effect, however, causes less deshielding than diamagnetic anisotropy.

Similar explanation can be used to explain the differences in carbon chemical shifts observed in the ^{13}C NMR for the CSO and SCSO. In the ^{13}C NMR spectra, both the CSO and SCSO have one signal each at around 14.1 ppm (methyl group at the end of the acyl chains in glyceride moiety). It is well separated from other signals and, hence, easily recognized. The same vaIues have been reported in literature (Gunstone, 2004, Okiemen et al., 2005, Sega et al., 2010). In the ^{13}C spectrum, the signals associated with the olefinic carbons appear highly deshielded at δ 127.09 to 131.85. ppm due to the diamagnetic anisotropic effect of the π system. Upon sulphonation, these signals were absent due to loss of the double bonds. The new signals which appeared at 52 and 72 ppm belong to the sp^3 hybridized carbons (C-S and C-O) formed after the sulphonation reactions. The slightly deshielded position of these signals is also due to the influence of the inductive effect of the electronegative sulphur and oxygen atoms.

Strength Properties of Leather Samples

The strength properties of the leather samples were shown in Figures 4, 5 and also in Table 6. The strength properties of the leather samples were determined both perpendicular and parallel to the backbone and the average results obtained. It is obvious from Figure 4 that despite the other processing techniques which involved conditioning, staking and inclusion of other additives to give leather a soft handle, both the SCSO and its blend with 7.5% castor oil demonstrated a clear improvement on leather without any fatliquor (NC), and was comparable to similar leather made with commercially available fatliquor (PC).

Castor seed oil is normally used in the tanning industry as a source of lubrication because of its humectant property. The strength of grain surface test results is shown in Figure 5. The lowest values for grain crack and grain burst are seen in the negative control (having average grain crack and grain burst strengths of 320 and 335 N respectively). The PC, Pure and Blend, had higher average grain crack of 400, 410 and 405 N respectively, and also grain burst strengths of 450, 415 and 425 N respectively. Proper lubrication of the leather fibres in PC, Pure and Blend brought about the increase in average grain crack and grain burst strength.

Strength properties have been given the greatest consideration in evaluating fatliquored leather because they give an indication of fibre lubricity (Waite, 1999; Burgess, 1993). The sulphonated portion of the oil interacts with the active centres in the collagen molecules of leather fibres and leaves the oil between the fibres. The elongation at break characterizes the softness, flexibility, strength and toughness of the leather matrix (Alexander et al., 1993). It is also evident from the strength properties results that the leather fatliquored using SCSO compared favourably with the commercial fatliquor.

Table 6
Other strength properties results

Properties	NC	PC	PURE	BLEND
Average Tear Load (Double Edge Tear) (N)	354.3	543.4	437.8	496.4
Average Tensile Strength (N/mm^2)	17.38	24.86	22.54	27.07
Average Elongation at Break (%)	28.48	38.04	35.88	37.14

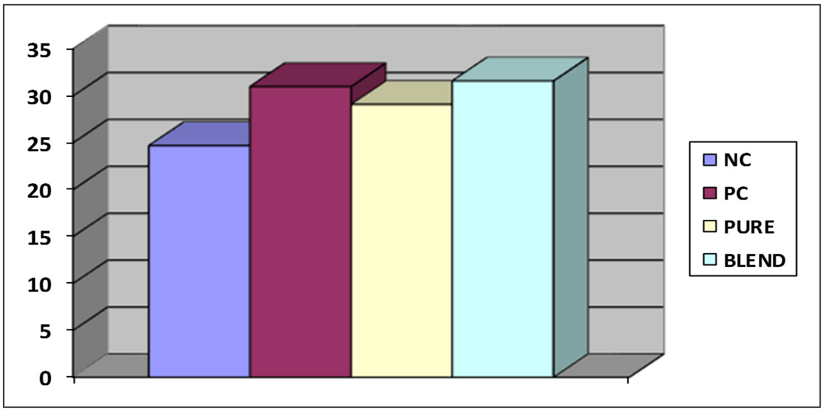


Figure 4. Softness test results

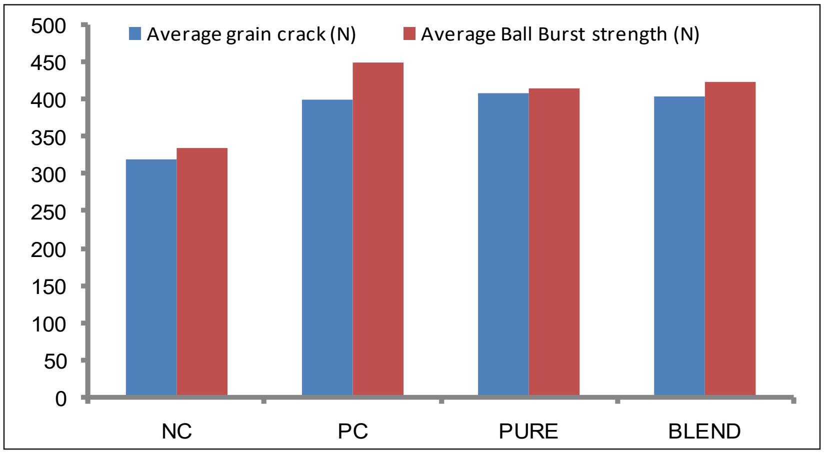


Figure 5. Strength of grain surface of leather

Sudan Stain Test Results

Staining renders visible, components within the section which would not be seen easily. Here, leather fibrils were soaked in Sudan (IV) dyes which were specific for fats and oils and are specially adapted to suit leather microscopy (Bugby, 1999). The presence of orange/red stains depicts the presence of fatliquors within the cross-sections of the leather samples B, C and D in contrast to A which was processed without fatliquor (Figure 6). Sample C also showed a high level of penetration of the prepared sulphonated oil within the leather. SCSO obviously compared favourably with the commercial fatliquor – Positive Control.

Scanning Electron Microscopy (SEM) Results

Figure 7 shows the surface morphology of the leather samples. From the SEM analysis results, the leather samples fatliquored with commercial fatliquor (positive control), prepared sulphonated fatliquor and its blend with 7.5% raw castor oil exhibited well opened up structures. In contrast, the leather which was not fatliquored had a split up structure but the fibres restuck after drying. The well opened up fibre structure depicts good lubrication.

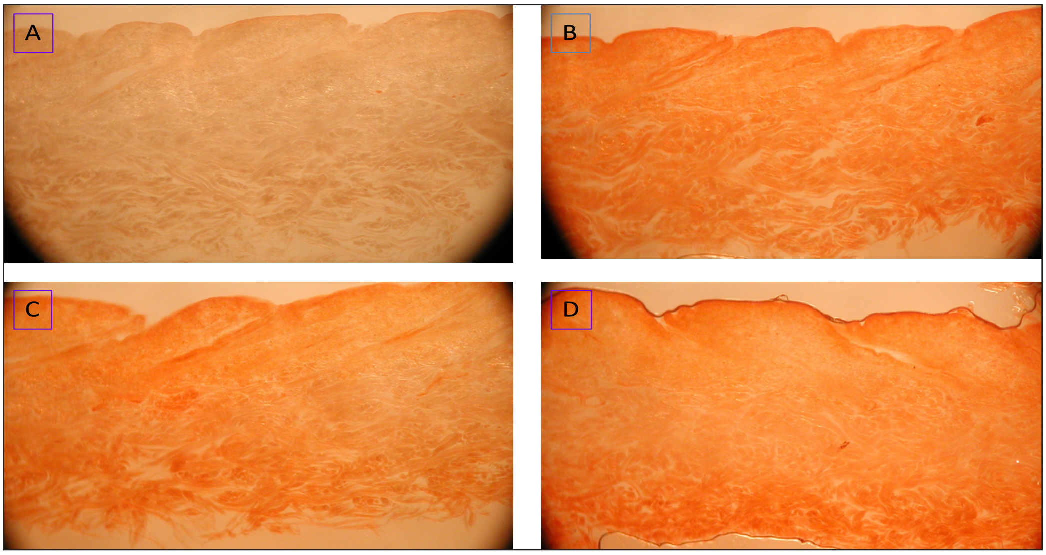


Figure 6. Staining Test Results Showing Cross-Section of Chrome Tanned Goatskins Processed with: **A**-NC; **B**- PC; **C**- A1 (PURE); **D**- A2 (BLEND)

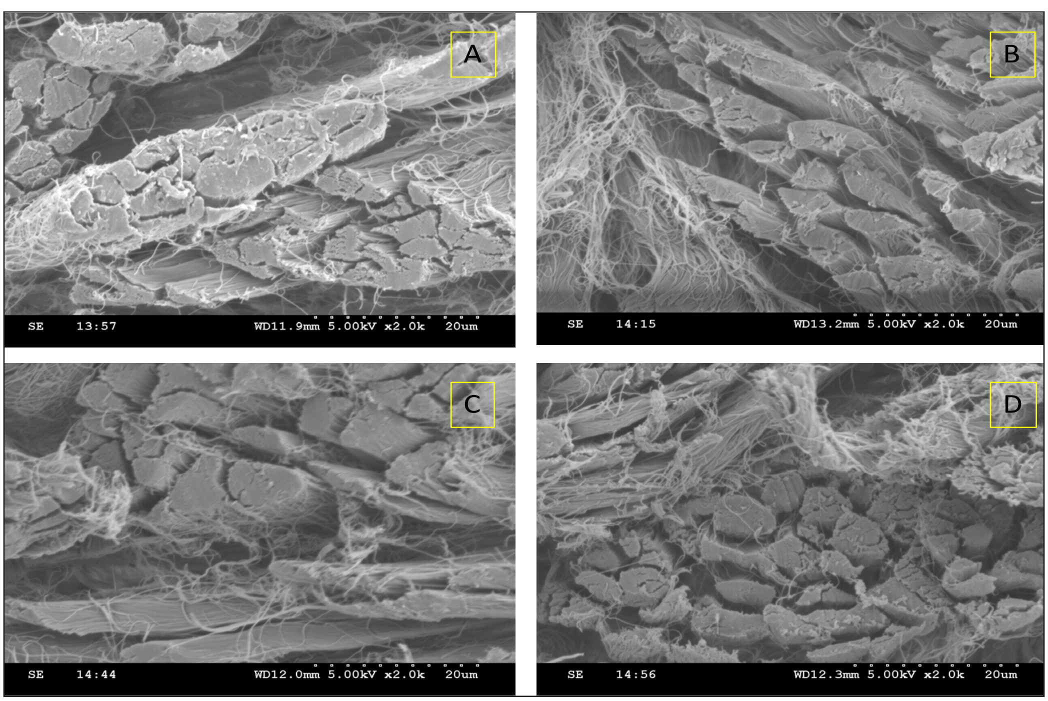


Figure 7. Scanning Electron Microscopy (X2000) of Chrome Tanned Goatskins Processed with: **A**-NC; **B**- PC; **C**- A1 (PURE); **D**- A2 (BLEND)

CONCLUSION

Structural characterizations performed by FT-IR, ^{1}H NMR and ^{13}C NMR analysis confirmed the sulphonation of CSO. It was observed from DSC results that both the CSO and SCSO were relatively thermally stable at the temperatures studied and would not likely decompose when being processed or in use as leather. The leather processed by the sulphonated *C. schweinfurthii* fatliquor had a comparable tensile strength, double edge tear, and grain strength with commercial/imported fatliquor. The stain test result showed the full lubrication of the *C. schweinfurthii* fatliquored leather and is comparable with leather from commercial fatliquor. This indicates that the sulphonated *C. schweinfurthii* fatliquor could be a rival with commercial products for the production of leather shoe upper.

ACKNOWLEDGEMENTS

The authors thank the management and staff of the Institute for Creative Leather Technologies (ICLT), University of Northampton, (UoN) Northampton, United Kingdom, for the support in terms of facilities, equipment and bench space offered for the work.

REFERENCES

Abayeh, O. J., Abdulrazaq, A. K., & Olaogun, R. (1999). Quality characteristics of *Canarium schweinfurthii*. *Plant Foods for Human Nutrition, 54*(1), 43-48.

Adewuyi, A., Awolade, P. O., & Oderinde, R. A. (2014). *Hura crepitans* seed oil: Alternative feedstock for biodiesel production. *Journal of Fuels, 2014*, 1-9.

Alexander, K. T. W., Corning, D. R., Covington, A. D., Stanley, A. M., & Stosic, R. G. (1993). The production of soft leather part 2: Drying and stress softening. *Journal of American Leather Chemists Association, 88*(7), 254-269.

Berg, M. J., Tymoczkko, J. L., & Styrer, L. (2002). *Biochemistry* (5th Ed.). New York, USA: W.H. Freeman Publishing.

BS 1723. (2015). *Leather, physical and mechanical tests: Determination of softness*. London, UK: British Standards Institution.

BS 2418. (2002). *Chemical, physical and mechanical and fastness tests: Sampling location*. London, UK: British Standards Institution.

BS 2419. (2002). *Chemical, physical and mechanical and fastness tests: Conditioning*. London, UK: British Standards Institution.

BS 3376. (2011). *Leather, physical and mechanical tests: Determination of tensile strength and percentage extension*. London, UK: British Standards Institution.

BS 3377. (2011). *Leather, physical and mechanical tests: Determination of tear load. Double edge tear*. London, UK: British Standards Institution.

BS 3379. (2015). *Leather, physical and mechanical tests: Determination of distension and strength of surface* (Ball Burst method). London, UK: British Standards Institution.

Bugby, A. (1999). Microscopy. In Leafe, M. K. (Ed.), *Leather technologists pocket book* (pp. 311-337). Withernsea, UK: Society of Leather Technologists and Chemists.

Burgess, D. (1993). General aspects of fatliquoring: An introduction to the application and chemistry of fat liquoring. *Journal of the Society of Leather Technologists and Chemists, 78*(2), 39-43.

Covington, A. D. (2011). *Tanning chemistry: The science of leather*. Cambridge, UK: Royal Society of Chemistry Publishing.

Daniels, R. P. (2001). Fatliquoring and softening agents. *World Leather, 14*, 68-70.

Evans, J. (2004). *Plantation forestry in the tropics* (2nd Ed.). London, UK: Oxford University Press.

Fashina, O. O., Craig-Schmidt, M., Colley, Z., & Hallman, H. (2008). Predicting melting characteristics of vegetable oils from fatty acid compositions. *Food Science Technology, 41*(8), 1501-1505.

Firestone, D. (1998). *Official methods and recommended practices of the American oil Chemists society* (15th Ed.). Champaign-Illinois, USA: American Oil Chemists Society.

Gunstone, F. D. (2004). *The chemistry of oils and fats: Sources, composition, properties and uses*. Oxford, UK: Blackwell Publishing Ltd.

Habib, M. A., & Alshammari, A. G. (2014). Leather fatliquor from hide fleshing. *Journal of the Society of Leather Technologists and Chemists, 98*(5), 199-203.

Heidemann, E. (1993). *Fundamentals of leather manufacturing*. Darmstadt, Germany: Eduard Roether KG Publishing.

ICLT SR15/31. (2015). *Leather manufacture: Shoe upper-chromium (non-compact).* Manual for Leather Processing. Northampton, UK: Institute for Creative Leather Technologies.

Janakiram, C. R. (2007). The theory and practice of fatliquoring. *World Leather, January edition*, 38-43.

Kaisersberger, E. (1989). DSC investigations of the thermal characterization of edible fats and oils. *Thermochim Acta*, *151,* 83-90.

Okieimen, F. E., Pavithran, C. & Bakare, I. O. (2005). Epoxidation and hydroxlation of rubber seed oil: One-pot multi-step reactions. *European Journal of Lipid Science and Technology, 107*(5), 330-336.

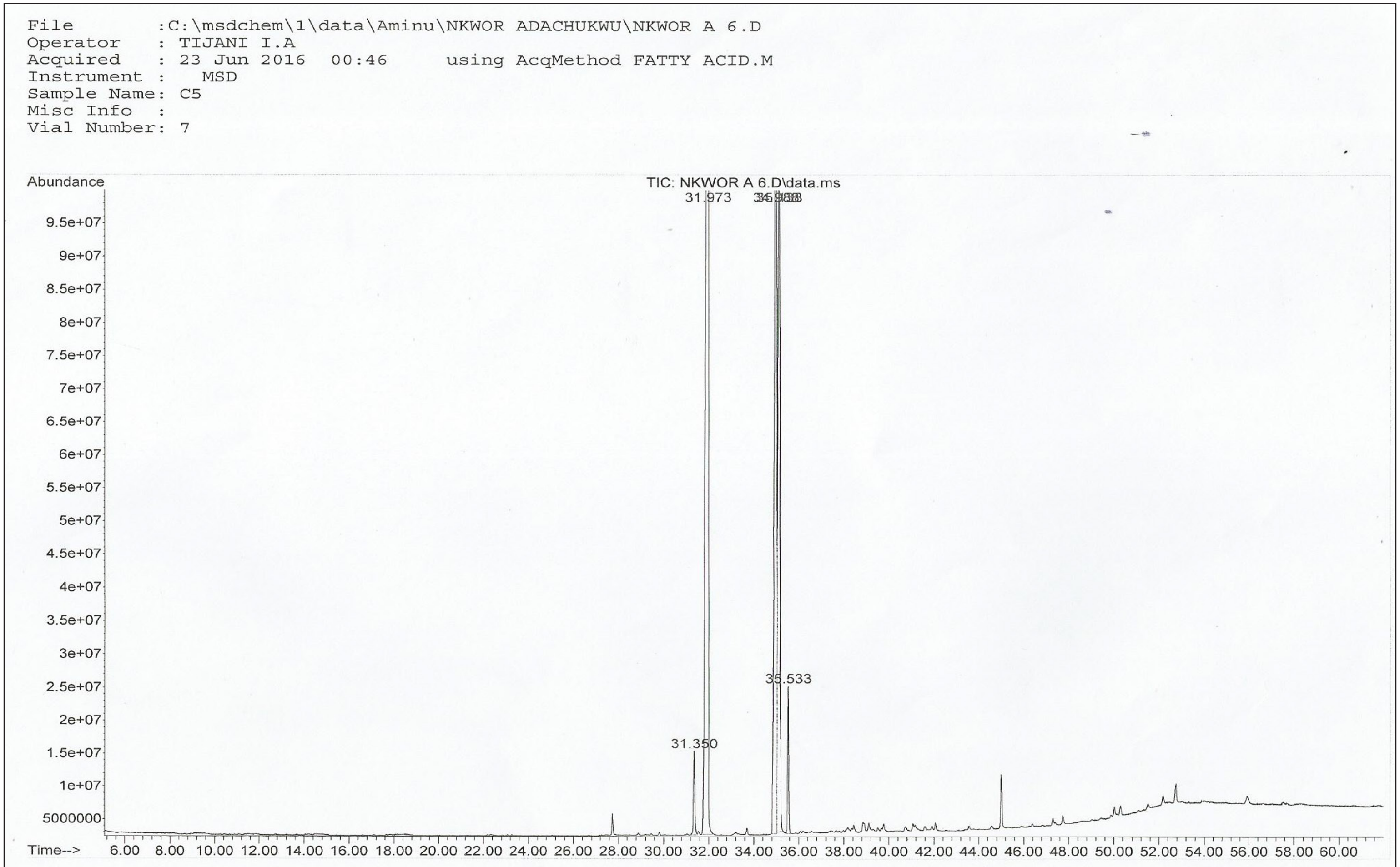
Sega, A., Zanardi, I., Chiasserini, L., Gabbrielli, A., Bocci, V., & Travagli, V. (2010). Properties of sesame oil by detailed ^{1}H and ^{13}C NMR assignments before and after ozonation and their correlation with iodine value, peroxide value and viscosity measurements. *Chemistry and Physics of Lipids, 163*(2), 148-156.

Steik, K. T. (1943). *U.S. Patent No. 2,328,931*. Washington, DC: U.S. Patent and Trademark Office.

Waite, T. (1999). Oils, fats, waxes and fatliquors. In Leafe, M. K., (Ed.), *Leather technologists pocket book* (pp. 125-148). Withernsea, UK: Society of Leather Technologists and Chemists.

APPENDIX A

Gas chromatogram of CSO methyl ester



APPENDIX B

Area percent report of CSO methyl ester

```
CHEMICAL ENG`G LAB.UNILORIN.     Area Percent Report

Data Path : C:\msdchem\1\data\Aminu\NKWOR ADACHUKWU\
Data File : NKWOR A 6.D
Acq On    : 23 Jun 2016  00:46
Operator  : TIJANI I.A
Sample    : C5
Misc      :
ALS Vial  : 7   Sample Multiplier: 1

Integration Parameters: rteint.p
Integrator: RTE
Smoothing : ON                          Filtering: 5
Sampling  : 1                            Min Area: 3 % of largest Peak
Start Thrs: 0.2                         Max Peaks: 100
Stop Thrs : 0                       Peak Location: TOP

If leading or trailing edge < 100 prefer < Baseline drop else tangent >
Peak separation: 5

Method    : C:\Users\admin\Desktop\METHODS\ESSENTIAL OILS_SCAN2.M
Title     :

Signal     : TIC: NKWOR A 6.D\data.ms
```

peak #	R.T. min	first scan	max scan	last scan	PK TY	peak height	corr. area	corr. % max.	% of total
1	31.350	4699	4722	4746	rBV2	12743186	53852393	4.06%	1.699%
2	31.973	4783	4834	4866	rBV4	149858768	1325676903	100.00%	41.812%
3	34.988	5338	5376	5385	rBV6	124954707	953797339	71.95%	30.083%
4	35.138	5385	5403	5432	rVB2	142093213	752829976	56.79%	23.745%
5	35.533	5458	5474	5503	rVB2	22362273	84379940	6.37%	2.661%

```
Sum of corrected areas:  3170536551

SSENTIAL OILS_SCAN2.M Fri Jun 24 08:57:10 2016
```

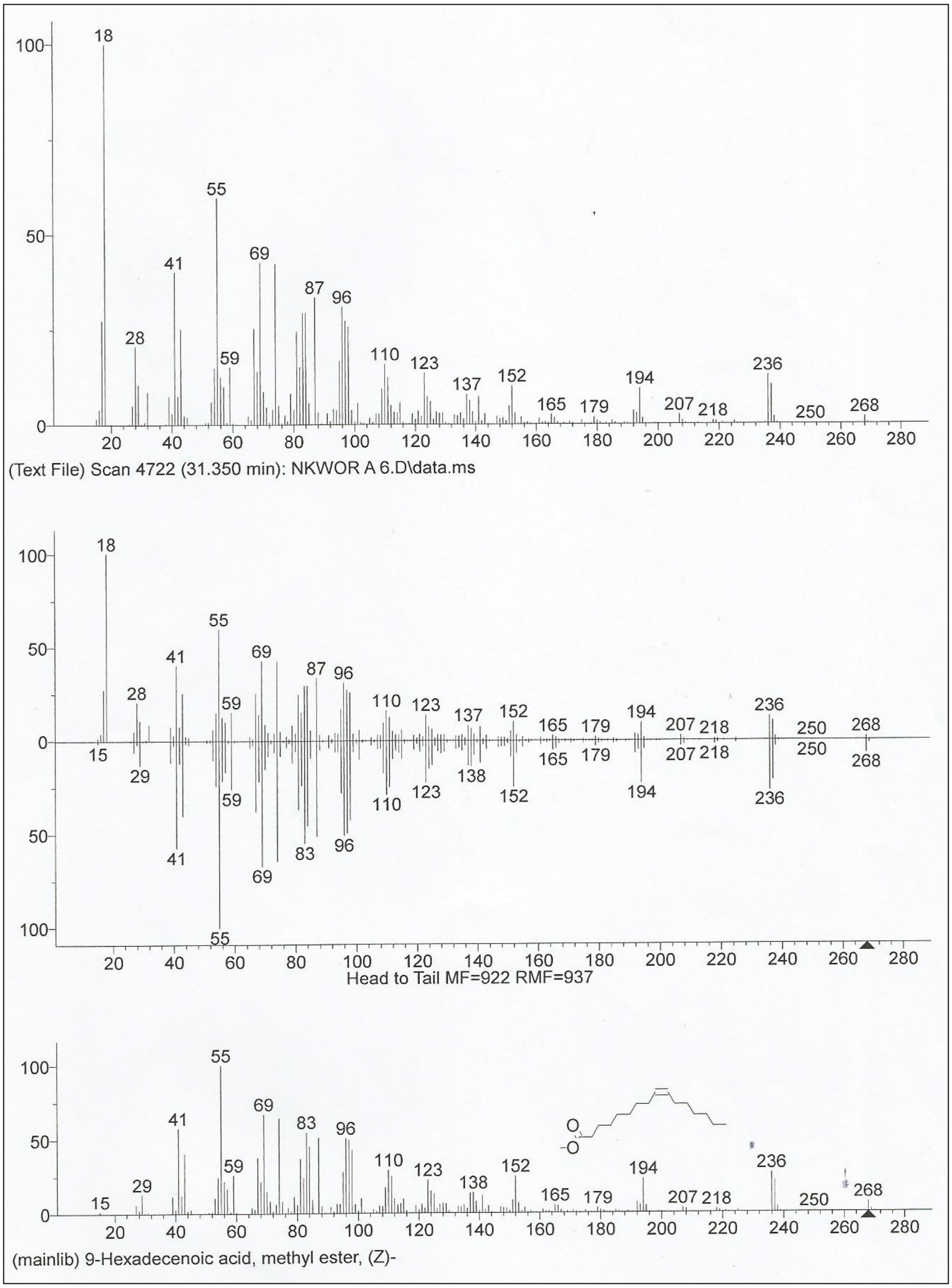
APPENDIX C

Library search report of CSO methyl ester

```
                                   CHEMICAL ENG`G LAB.UNILORIN.     Library Search Report

  Data Path : C:\msdchem\1\data\Aminu\NKWOR ADACHUKWU\
  Data File : NKWOR A 6.D
  Acq On    : 23 Jun 2016  00:46
  Operator  : TIJANI I.A
  Sample    : C5
  Misc      :
  ALS Vial  : 7   Sample Multiplier: 1

  Search Libraries:   C:\DATABASE\demo.l                Minimum Quality:  15
                      C:\Database\NIST11.L              Minimum Quality:  90

  Unknown Spectrum:   Apex minus start of peak
  Integration Events: RTE Integrator - lscint.e

Pk#     RT  Area%          Library/ID                 Ref#     CAS#   Qual
_______________________________________________________________________________
  1  31.350  1.70 C:\Database\NIST11.L
                 7-Hexadecenoic acid, methyl ester, 117506 056875-67-3 99
                  (Z)-
                 9-Hexadecenoic acid, methyl ester, 117513 001120-25-8 98
                  (Z)-
                 9-Hexadecenoic acid, methyl ester, 117511 001120-25-8 98
                  (Z)-

  2  31.973 41.81 C:\DATABASE\demo.l
                 Methyl palmitate                         4 000112-39-0 70

  3  34.988 30.08 C:\Database\NIST11.L
                 9,12-Octadecadienoic acid (Z,Z)-,  139724 000112-63-0 99
                 methyl ester
                 9,12-Octadecadienoic acid, methyl  139708 002462-85-3 99
                 ester
                 9,12-Octadecadienoic acid, methyl  139733 002566-97-4 98
                 ester, (E,E)-

  4  35.138 23.74 C:\Database\NIST11.L
                 14-Octadecenoic acid, methyl ester 141287 056554-48-4 99
                 11-Octadecenoic acid, methyl ester 141313 001937-63-9 98
                 , (Z)-
                 9-Octadecenoic acid, methyl ester, 141309 001937-62-8 97
                  (E)-

  5  35.533  2.66 C:\DATABASE\demo.l
                 Methyl palmitate                         4 000112-39-0 72


ESSENTIAL OILS_SCAN2.M Fri Jun 24 08:57:20 2016
```

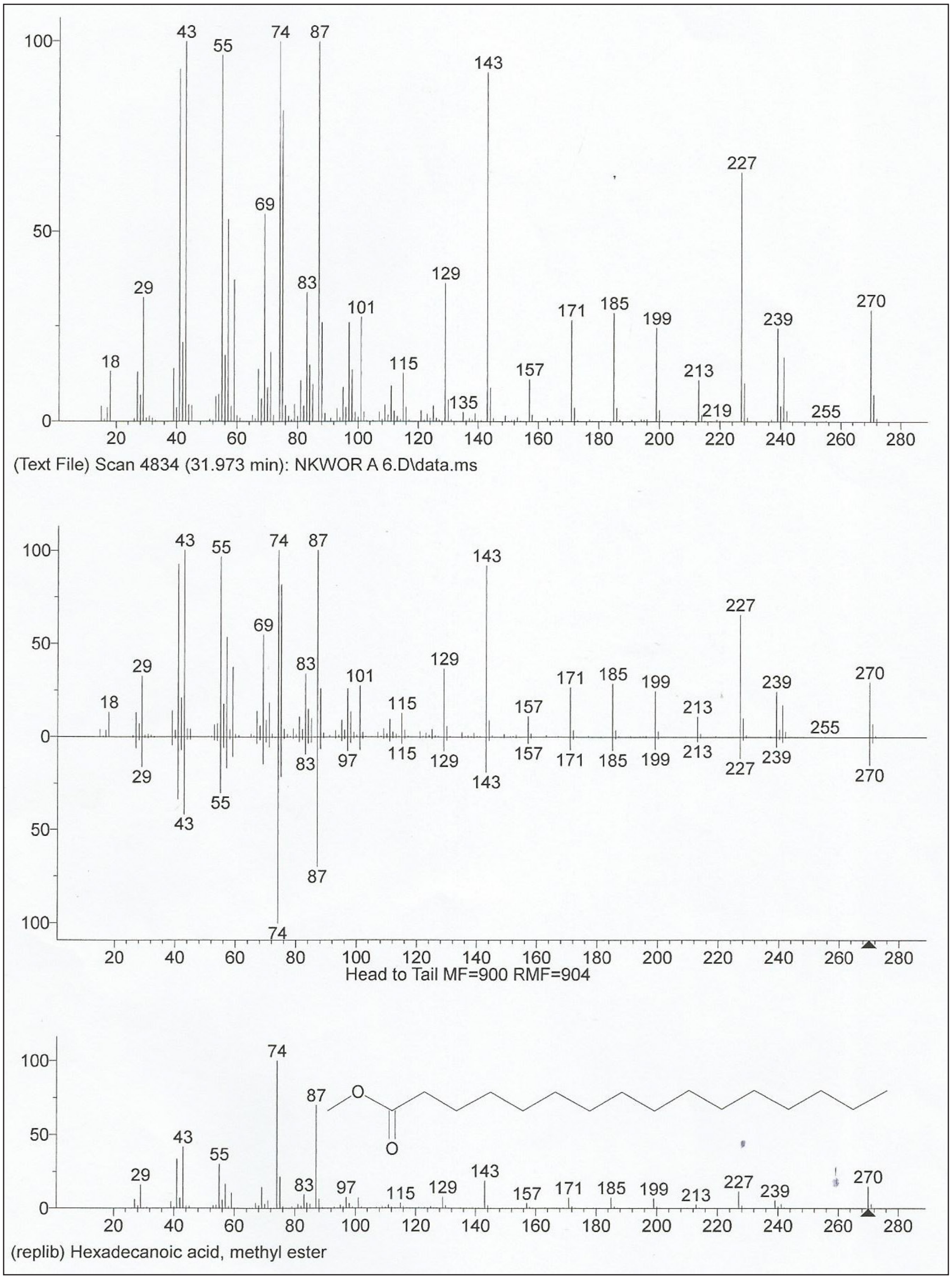
APPENDIX D

CSO methyl ester- mass spectrum for area percent (1.699 %, 31.350 min)



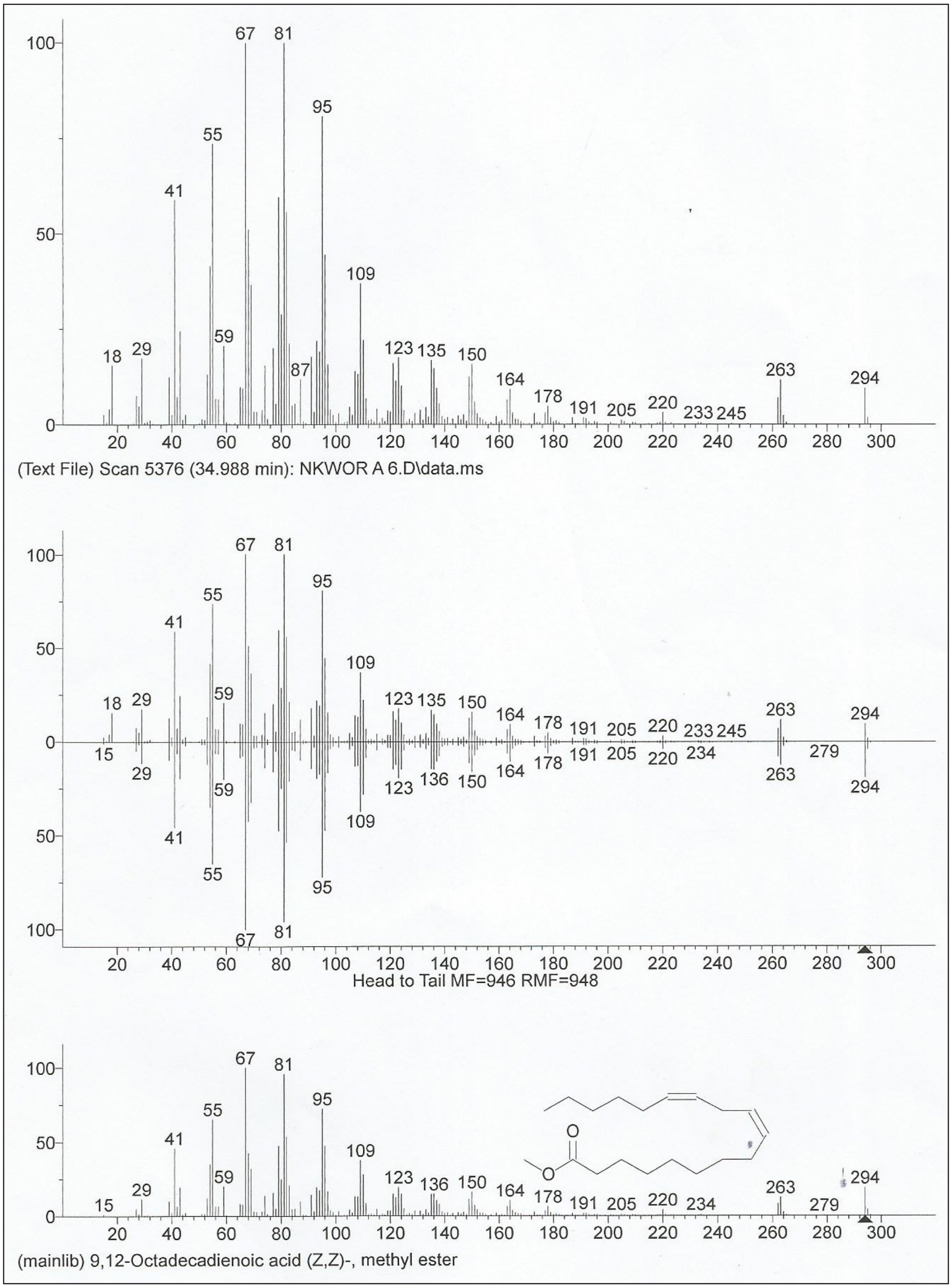
APPENDIX E

CSO methyl ester- mass spectrum for area percent (41.81 %, 31.973 min)



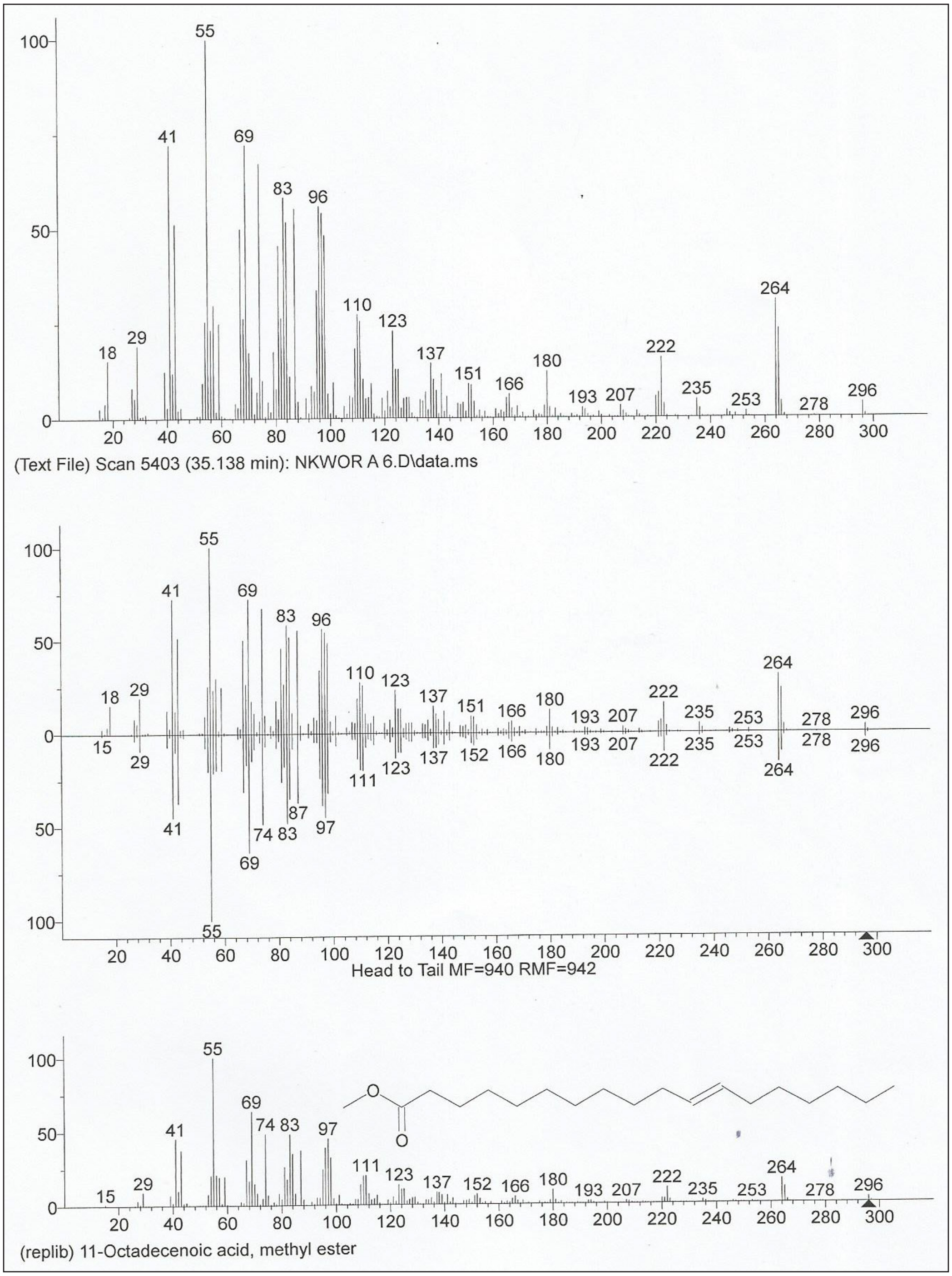
APPENDIX F

CSO methyl ester- mass spectrum for area percent (30.08 %, 34.988 min)



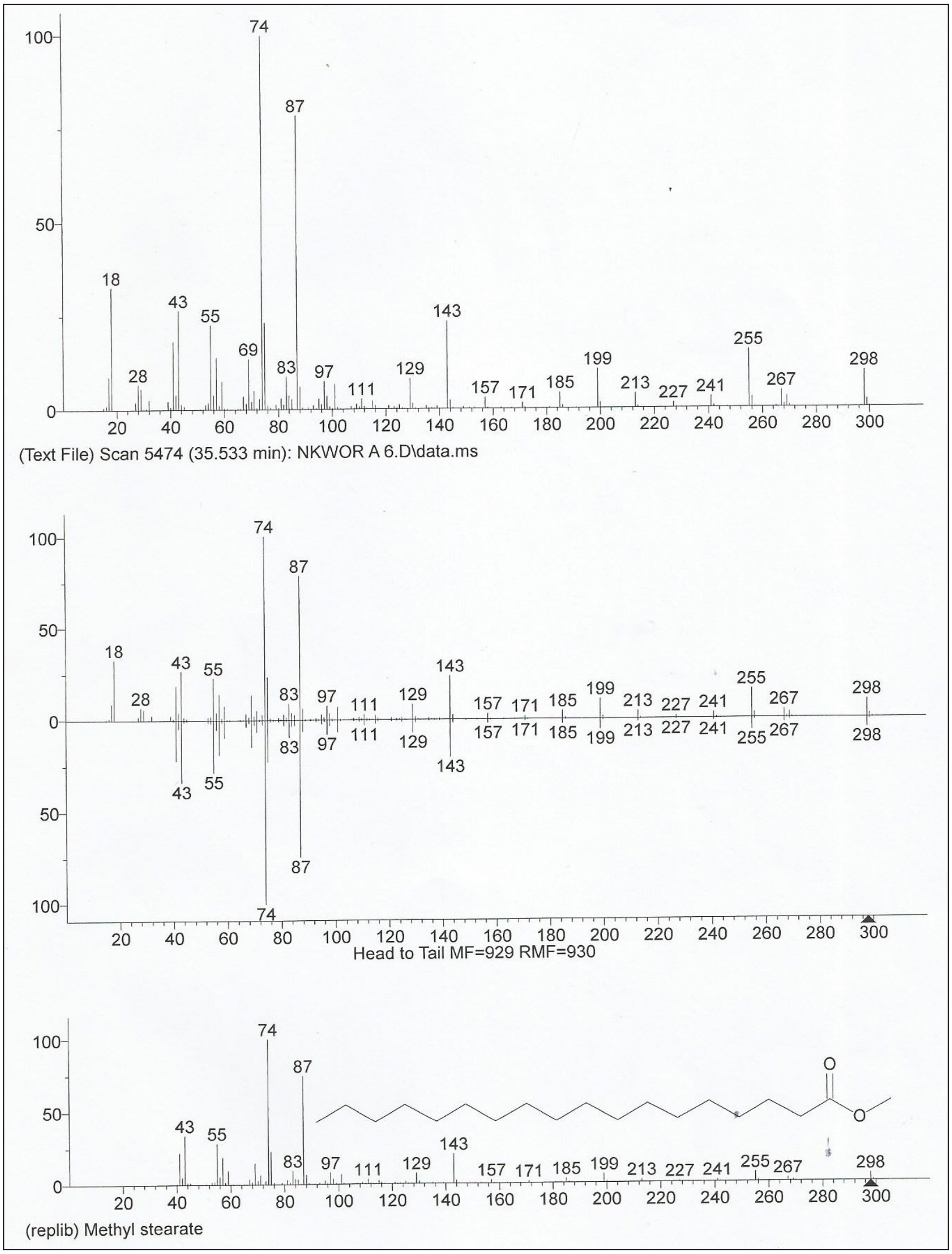
APPENDIX G

CSO methyl ester- mass spectrum for area percent (23.74 %, 35.138min)



APPENDIX H

CSO methyl ester- mass spectrum for area percent (2.66 %, 35.533 min)



Pertanika J. Sci. & Technol. 27 (4): 2245 - 2263 (2019)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Assessment of Runoff via Precipitation using Neural Networks: Watershed Modelling for Developing Environment in Arid Region

Sandeep Samantaray, Abinash Sahoo and Dillip Kumar Ghose*

Department of Civil Engineering, National Institute of Technology Silchar, Assam, 788010 India

ABSTRACT

This work describes the application of three different neural network, (i) Back propagation neural network (BPNN), (ii) Layer Recurrent Neural Network (LRNN) and (iii) Radial Basis Fewer Network (RBFN) model, to predict runoff. Here, two scenarios were considered for developing the models. Scenario 1 exclusive of evapotranspiration and Scenario 2 with evapotranspiration are considered for experiencing the impact on runoff. Performance indicators entailed Scenario 2 performed best as compared to Scenario 1. Two watersheds Loisingha, and Saintala were considered for study. In Loisingha watershed, LRNN performed best with architecture 4-3-1 following tangential sigmoid transfer function. At Saintala, both LRNN and BPNN performed in parallel with small deviation of prediction and LRNN performed best among three networks with model architecture 4-2-1 using Log-sig transfer function for predicting runoff.

Keywords: Evapotranspiration, neural network, precipitation, runoff, temperature, watershed

INTRODUCTION

Water availability is the computation of runoff resulting from precipitation and its abstracts on watershed. The correlation between precipitation and evapotranspiration involves mapping for measuring runoff. Inclusion of temperature in addition to the above inputs is helpful to measure the accuracy of model. In this context, the present study gives an idea for estimating runoff. Estimation of runoff is achieved using hydrological modelling to understand catchment behaviours, impacts of climate, and land use changes. Common methods for estimation include the physical, statistical, and combined approaches to analyze available data.

ARTICLE INFO

Article history:
Received: 04 October 2018
Accepted: 04 June 2019
Published: 21 October 2019

E-mail addresses:
samantaraysandeep963@yahoo.com (Sandeep Samantaray)
bablusahoo1992@gmail.com (Abinash Sahoo)
dillipghose2002@gmail.com (Dillip Kumar Ghose)
* Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

Artificial Neural Network (ANN) was utilized for computing the model using daily stream flows over Leaf River Basin (Smith & Eli, 1995; Tokar & Johnson 1999). Ghose and Samantaray (2018a) used regression and BPNN technique for hydrological modelling at Suktel watershed. Lungu (1991) observed the non-linear relationship between monthly rainfall and runoff. Srinivasulu and Jain (2006) found the factors which affected the runoff response of a watershed. Chen and Adams (2006) found conceptual rainfall-runoff models that were broadly used in hydrological modelling. Ghose and Samantaray, (2018b) employed BPNN and RBNN techniques to evaluate water table instability in sparse rainfall province, Odisha. Zhao (1992) implemented Xinanjiang (XAJ) model for measuring rainfall-runoff in China. Kothyari and Singh (1999) utilised different methods for estimating monthly runoff during monsoon months from June to October, in Northwest Iran. RBFN showed similar modelling results to that of multilayer perceptron neural network (NN) in building relationship between rainfall-runoff (Dawson & Wilby, 2001; Kumar et al., 2005; Shamseldin et al., 2007; Zakermoshfegh et al., 2008). Growth of hybrid NNs with conceptual model has received substantial consideration (Lee et al., 2002; Jain & Srinivasulu, 2006). Ghose and Samantaray, (2018c) employed BPNN, RBFN and Non Linear Multilayer Regression techniques to develop sediment rating models for period of monsoon season for Mahanadi river basin upstream of Tikarapada. Kisi et al. (2013) compared the accuracy of Gene Expression Programming, Adaptive Neuro-Fuzzy Inference System and ANN techniques in modelling rainfall-runoff process in Turkey. Chandwani et al. (2015) presented applications in modelling rainfall-runoff relations which replaced time consuming conventional mathematical techniques. Asadi et al. (2013) proposed a model which was a hybrid of data pre-processing methods, genetic algorithms and Levenberg–Marquardt (LM) algorithm for learning feed forward neural networks. Javan et al. (2015) investigated impacts of climate change on runoff using ANN for various gauging station of Gharehsoo River Basin in the northwest of Iran. Shoaib et al. (2014) employed hybrid Multilayer Perceptron Neural Network (MLPNN) and the Radial Basis Function Neural Network (RBFNN) to predict rainfall-runoff modelling at Brosna catchment. Descent neural network was considered to predict sediment load at Salabheta gauging station, India (Samantaray & Ghose, 2018). The objective is to develop correlation of rainfall-runoff and prediction of runoff through various neural networks considering precipitation, maximum and minimum temperature as input. Also addition of evapotranspiration with scenario one is employed to establish rainfall runoff co-relation.

STUDY AREA AND DATA COLLECTION

Two watersheds of Bolangir, Odisha, India, were taken into consideration for the proposed study area. The watersheds are located at Loisingha and Saintala having geographical area 317.6 Sq km, and 454.43 Sq km as shown in Figure 1. The purpose of this study was to predict runoff in two watersheds to assess the drainage capacity of watershed during monsoon period ranging from 1998 to 2017. The two watersheds are situated in the upstream of Hirakund reservoir. The geo coordinate of the watersheds are latitude 21º88′ 36″ N and longitude 84º 90′ 75″ E.

The average monthly precipitation, highest monthly average temperature and least monthly average temperature and evapotranspiration data for month of monsoon (May to October) from the period 1998-2017 spanning over 20 years were collected from India Meteorological Department (IMD) Bhubaneswar. The monthly runoff data were collected from soil conservation office, Bolangir. Average rainfall is within the range of 130cm to 141cm. Since the study area is within less rainfall region, prediction of runoff is necessary.

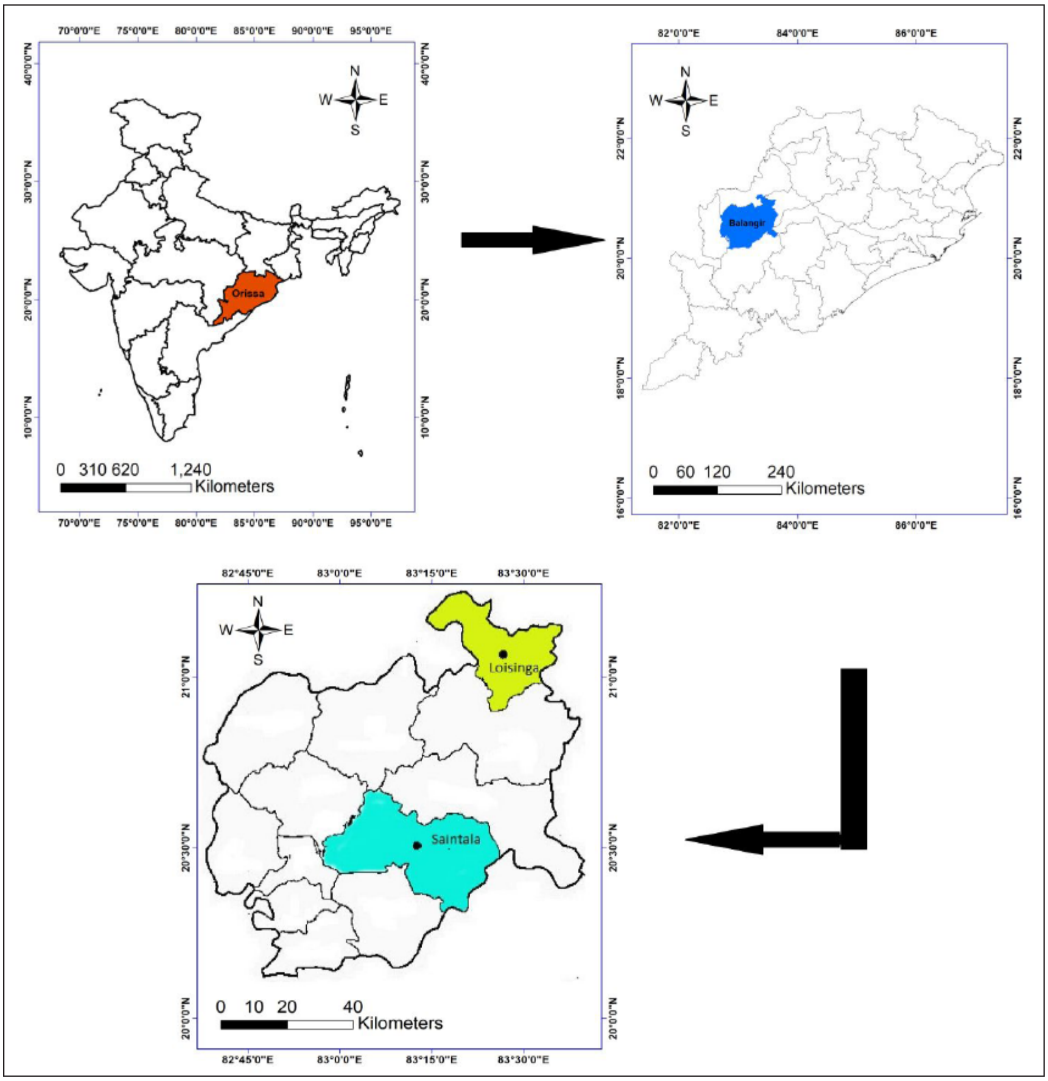


Figure 1. Study area

MATERIAL AND METHOD

Back-propagation Neural Network

The multi-layer perceptron network with supervised learning model which minimizes the error using weight adoption by applying back propagation of error is identified as BPNN. Back propagation was used to compute the gradient of the error of the network with respect to adjustable weights of the network. Further, this gradient used a modest stochastic gradient descent algorithm to determine weights that could reduce the error. Back propagation permits quick convergence on acceptable local minima for error in the kind of networks to which it is suited. It is significant to point that back propagation networks are essentially multiple layers (with single input, single hidden and single output layer). For the hidden layer to serve any suitable function, multiple layer networks should have non-linear activation functions to trigger the multiple layers. Back propagation neural network is a three-layered feed forward architecture, which accelerates input layer, hidden layer and output layer for functioning of back propagation activities in three stages, i.e. learning or training, testing or inferences and validation. Figure 2 shows the l-m-n (l input neurons, m hidden neurons, and n output neurons) architecture of a back propagation neural network model.

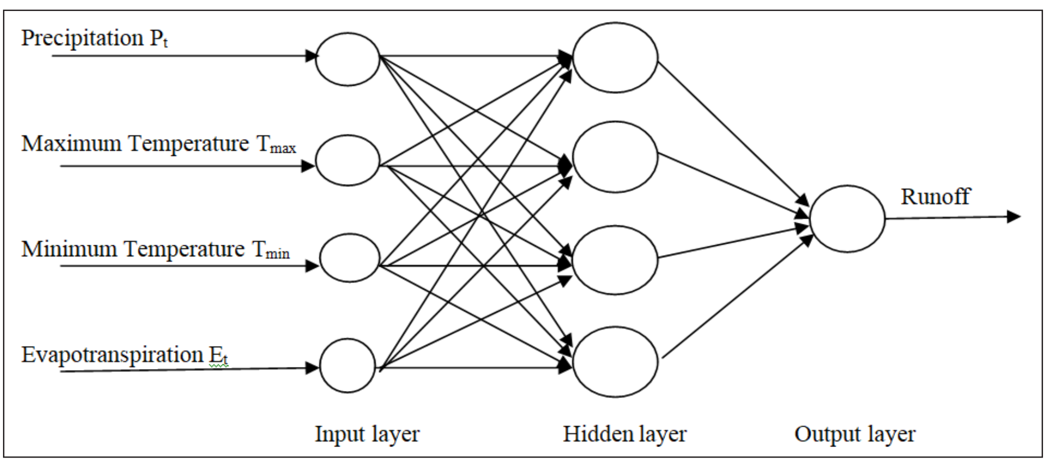


Figure 2. BPNN model with 4 input

Radial Basis Function Network

Basically RBFN comprises a huge number of simple and highly interconnected artificial neurons and can be organized into numerous layers, i.e. input layer, hidden layer, and output layer as presented in Figure 3.

Algorithm of RBFN (Fahrmeir et al., 2013). Following steps shows the basic algorithm of RBFN.

Input layer:

Inputs of normalized data is to be processed connecting with direct transfer function and pattern of interaction is to be directed through weights of neurons. Number of nodes in the input layer is equal to the number of input parameter.

Hidden layer:

The hidden layer helps in processing patterns merged through input and weights interlinked through input-output. Radial symmetry functions are used as transfer functions for processing the net. The output of unit i_m unit $V_i(x_i)$ in the hidden layer

$$V_i x_i = e^{-\frac{x^2}{\sigma^2}} \qquad (1)$$

Where $x^2 = (x_{ji} - \widehat{x_{ji}})^2 \ x_{ji}$

In which x_{ji} is the centre of RBF unit for input variables

σ is the width of RBF unit

$\widehat{x_{ji}}$ average of input variables

Output layer:

In this layer the output of the neural network is evaluated. The output of the neuron network is evaluated as

$$y_{net} = \sum_{i=1}^{H} w_{im} v_i(x_i) + w_0 \quad y_{net} \ w_{im} \ w_0 \qquad (2)$$

Where H = number of hidden node

y_{net} = output value of m_{th} node in output layer for the n_{th} incoming pattern

w_{im} = weight between i_{th} RBF unit and m_{th} output node

w_0 = biasing term at n_{th} output node

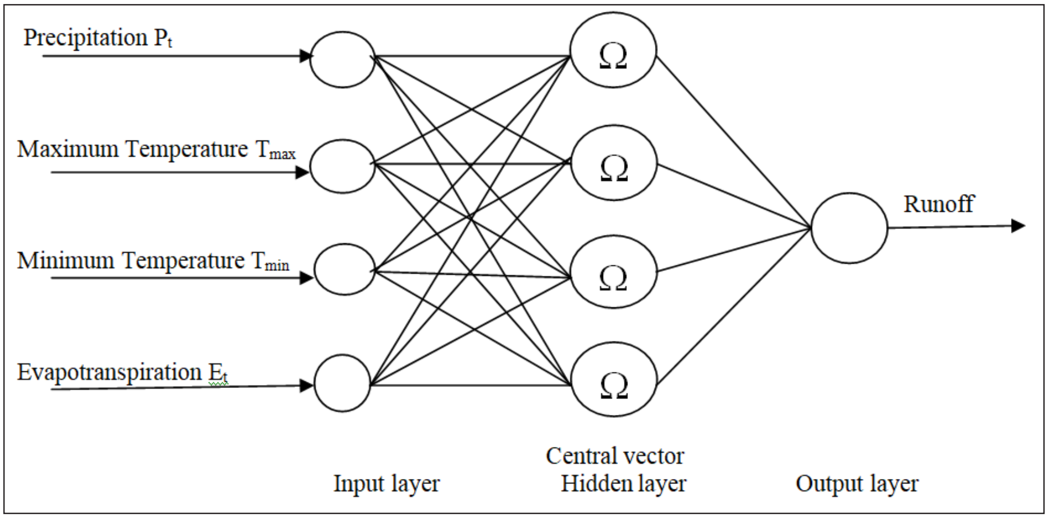


Figure 3. Architecture of RBFN.

Layer Recurrent Neural Network

LRNN is developed through an Elman network which is a three-layer network with the addition of context units. The hidden layer is connected through context units merged with fixed unit weight. The fixed back-connections memorize copy of the ancestor of the hidden units for further processing. The network is allowing sequence prediction with multilayer perceptron networks.

Elman network is defined as

$$H_t = \sigma_h(W_h X_t + U_h H_{t-1} + b_h) \tag{3}$$

$$Y_t = \sigma_y(W_y H_t + b_y) \tag{4}$$

Where x_t = input unit
H_t = Hidden unit
Y_t = output unit
W, U, b= parameter matrices
σ_y, σ_h = Activation function

Each input unit is linked to every one hidden unit, in combination with context unit shown in Figure 4. The recurring connections permit the hidden units to recycle the information over multiple time steps, thus letting to determine temporal information that is confined in the sequential input and relevant to the target function.

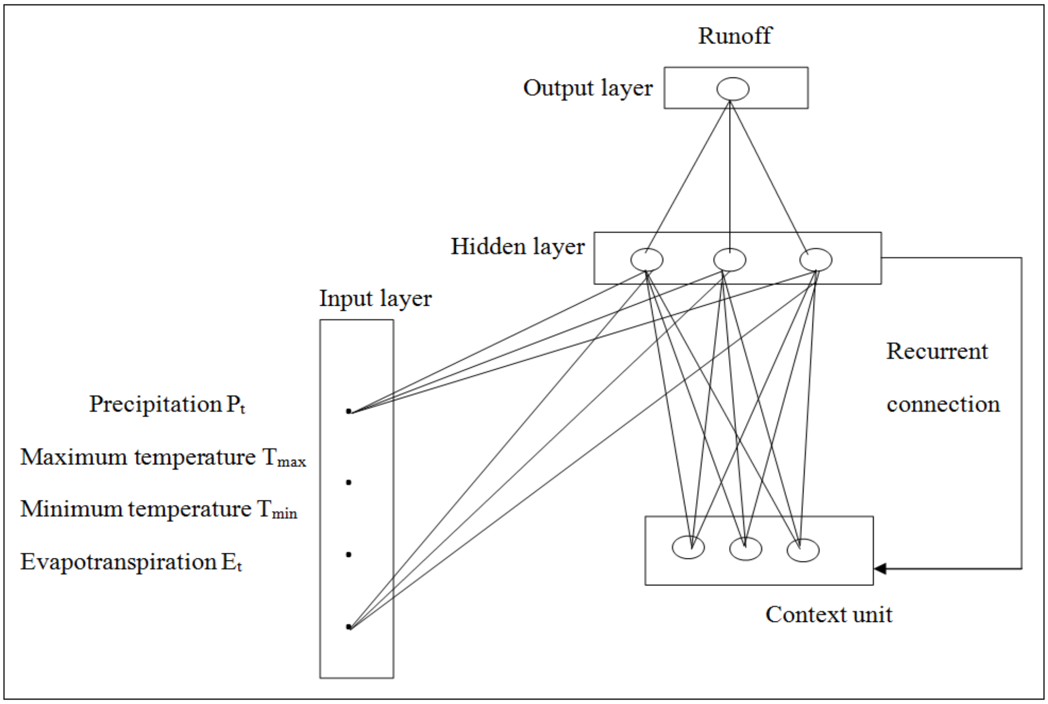


Figure 4. Architecture of LRNN

Processing and Preparation of Data

The monthly rainfall, highest monthly temperature, least monthly temperature were collected from IMD, Bhubaneswar and daily runoff data were collected from soil conservation office, Bolangir, Odisha, India for period of the monsoon months (May to October), from 1998-2017. The data collected from 1998-2011 (70% of data set) were used for training and from 2012-2017 (30%) were used for testing the network. Daily data were converted into monthly data, which finally helped in training and testing the model. The input and output data were scaled in such a way that every data fell within a quantified range before training. This process is known as normalization so that the normalized values are confined within the range of 0 to 1. The normalization equation which is used for scaling the data is

$$X_t = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{5}$$

Where X_t = transformed data series, X = original input data series, X_{min} = minimum of original input data series, X_{max} = maximum of original input data series.

Evaluating Criteria

Coefficient of determination, mean square error (MSE) and root mean square error (RMSE) are the evaluating standards to determine the best model. To select the perfect model for this area of study, the condition is MSE, RMSE must be minimum and coefficient of determination must be maximum.

Co-efficient of determination:

$$(\mathrm{R}^2) = 1 - \frac{(\sum_{\mathrm{i}=1}^{\mathrm{N}} x_{comp}^i - \bar{x}_{comp})^2}{(\sum_{\mathrm{i}=1}^{\mathrm{N}} x_{obs}^i - \bar{x}_{obs})^2} \tag{6}$$

The coefficient of determination value specifies the percent of variation in one variable explained by the other variable.

$$\text{MSE} = \frac{1}{n}\sum_{j=1}^{n}(x_{comp}^i - x_{obs}^i)^2 \tag{7}$$

$$\text{RMSE} = \frac{\sum_{i=1}^{N}(x_{comp}^i - \bar{x}_{comp})(x_{obs}^i - \bar{x}_{obs})}{\sqrt{\sum_{i=1}^{N}(x_{comp}^i - \bar{x}_{comp})^2}(x_{obs}^i - \bar{x}_{obs})^2} \tag{8}$$

Where i= dummy variable

N= Number of data

x_{comp}^i = Predicted data

x_{obs}^i = Observed data

$\bar{x}_{comp}$ =Mean predicted data

$\bar{x}_{obs}$ = Mean observed data

RMSE is the root of mean squared error. RMSE also has non negative values and close to zero gives the best performance.

RESULTS AND DISCUSSIONS

Analysis of Results with Scenario 1 and Scenario 2

Scenario 1 was estimated by considering precipitation, maximum temperature, and minimum temperature as input for predicting runoff in two watersheds. The basic equation used for predicting the runoff in two watersheds basing upon the equation $Q_{t} = f(P_{t}, T_{max}, T_{min},)$ for Scenario1. Similarly Scenario 2 was estimated by considering precipitation, maximum temperature, minimum temperature, and evapotranspiration as input for predicting runoff in two watersheds. Equation used for predicting the runoff in two watersheds basing upon the equation $Q_{t}= f(P_{t}, T_{max}, T_{min}, E_{t})$ for scenario 2.

Results for Scenario 1 at Loisingha. Scenario 1 was estimated by considering precipitation, maximum temperature, and minimum temperature as input for predicting runoff at Loisingha watershed. The results are presented in Table 1, Table 2, and Table 3.

Table 1
Results of BPNN for Scenario 1

Scenario 1	Transfer Function	Architec-ture	MSE		RMSE		R^2	
			Training	Testing	Training	Testing	Training	Testing
P_{t}, T_{max}, T_{min}	Tansig	**3-2-1**	**0.00883**	**0.03367**	**0.09396**	**0.18351**	**0.8691**	**0.8376**
		3-3-1	0.00608	0.00459	0.07823	0.06776	0.7809	0.7289
		3-4-1	0.00939	0.07341	0.09691	0.27093	0.7733	0.7194
		3-5-1	0.00724	0.64751	0.08509	0.80464	0.8004	0.7475
		3-9-1	0.00115	0.00324	0.03394	0.05726	0.7768	0.7067
	Logsig	3-2-1	0.00963	0.00865	0.02151	0.03432	0.8472	0.7989
		3-3-1	0.00719	0.00686	0.03453	0.06392	0.8591	0.8112
		3-4-1	0.00082	0.00893	0.02059	0.03595	0.8324	0.7817
		3-5-1	**0.00078**	**0.00727**	**0.02429**	**0.06358**	**0.8713**	**0.8215**
		3-9-1	0.00818	0.00637	0.03435	0.06131	0.8781	0.8165
	Purelin	3-2-1	0.00856	0.00367	0.02925	0.05802	0.8553	0.8015
		3-3-1	0.00991	0.00471	0.03148	0.06380	0.7961	0.7511
		3-4-1	0.00975	0.00179	0.03122	0.04448	0.8431	0.7927
		3-5-1	0.00972	0.00378	0.03117	0.06138	07928	0.7521
		3-9-1	**0.00198**	**0.00512**	**0.03461**	**0.07079**	**0.8693**	**0.8177**

Table 2
Results of LRNN network for Scenario 1

Scenario 1	Transfer Function	Architecture	MSE		RMSE		R^2	
			Training	Testing	Training	Testing	Training	Testing
P_t, T_{max}, T_{min}	Tansig	**3-2-1**	**0.00105**	**0.00206**	**0.03170**	**0.04538**	**0.9215**	**0.8889**
		3-3-1	0.00551	0.01735	0.07711	0.13171	0.9004	0.8672
		3-4-1	0.00963	0.03709	0.09813	0.19253	0.8841	0.8521
		3-5-1	0.00102	0.03021	0.03165	0.17381	0.8933	0.8662
		3-9-1	0.00187	0.00514	0.03587	0.07080	0.9137	0.8748
	Logsig	3-2-1	0.00653	0.00516	0.01151	0.00242	0.8772	0.8289
		3-3-1	0.00419	0.00386	0.02453	0.01392	0.8891	0.8412
		3-4-1	0.00062	0.00593	0.02059	0.01095	0.8624	0.8117
		3-5-1	0.00048	0.00427	0.01429	0.00348	0.8913	0.8515
		3-9-1	**0.00418**	**0.00337**	**0.01435**	**0.00131**	**0.9081**	**0.8665**
	Purelin	3-2-1	0.00549	0.00427	0.02458	0.05802	0.8295	0.7831
		3-3-1	0.00693	0.00491	0.02964	0.05721	0.8463	0.8174
		3-4-1	0.00629	0.00582	0.02991	0.04082	0.8417	0.8106
		3-5-1	**0.00294**	**0.00184**	**0.03092**	**0.05739**	**0.8735**	**0.8291**
		3-9-1	0.00458	0.00354	0.03071	0.06392	0.8432	0.8051

Table 3
Results of RBFN network for Scenario 1

Scenario 1	Architecture	MSE		RMSE		R^2	
		Training	Testing	Training	Testing	Training	Testing
P_t, T_{max}, T_{min}	4-0.2-1	0.00429	0.00628	0.00432	0.00418	0.8014	0.7528
	4-0.3-1	0.00372	0.08125	0.00201	0.02016	0.8273	0.7938
	4-0.5-1	0.00512	0.05934	0.02639	0.08429	0.8195	0.7682
	4-0.7-1	0.00384	0.01937	0.02732	0.06418	0.8216	0.7828
	4-0.9-1	**0.00502**	**0.00923**	**0.02013**	**0.04962**	**0.8526**	**0.8093**

Results for Scenario 2 at Loisingha. Here to evaluate the model efficiency of various architectures, different transfer functions like tangential sigmoidal, logarithmic sigmoidal and purelin were used to establish the model performance. Every architecture criteria for evaluation is mean square error training, testing, root mean square error training, testing and coefficient of determination. In Table 4 for Tan-sig function in BPNN, 4-2-1, 4-3-1, 4-4-1, 4-5-1, 4-9-1 architectures were taken into consideration for computation of performance. For Tan-sig function the best model architecture was found to be 4-9-1 which possesses MSE training value 0.00679, MSE testing value 0.00376, RMSE training value 0.02645 RMSE testing value 0.05404 and coefficient of determination for training 0.9587 and testing 0.9287. For Log-sig 4-2-1, 4-3-1, 4-4-1, 4-5-1, 4-9-1 architectures were taken into consideration for computation of performance. For Log-sig function the best model architecture was found to be 4-2-1 which possess MSE training value 0.00463 MSE testing value 0.00165, RMSE training value 0.01591 RMSE testing value 0.03432 and coefficient of determination value training 0.9662, testing value 0.9327. For Purelin 4-2-1, 4-3-1, 4-4-1, 4-5-1, 4-9-1 architectures were taken into consideration for computation of performance. For Purelin function the best model architecture was found to be 4-4-1 which possess MSE training value 0.00851 MSE testing value 0.00325, RMSE training value 0.02978, RMSE testing value 0.05374 and coefficient of determination value training 0.9592 and testing value 0.9218. The detailed results are presented in Table 4.

Table 4
Results of BPNN at Loisingha for Scenario 2

Scenario 2	Transfer Function	Architec-ture	MSE		RMSE		R^2	
			Training	Testing	Training	Testing	Training	Testing
P_t, T_{max}, T_{min}, E_t	Tansig	4-2-1	0.00806	0.00417	0.02839	0.06163	0.9412	0.9153
		4-3-1	0.00795	0.00575	0.02815	0.07584	0.9355	0.9073
		4-4-1	0.00472	0.00256	0.09175	0.05493	0.9075	0.8807
		4-5-1	0.00471	0.00761	0.02805	0.01732	0.9289	0.8991
		4-9-1	**0.00679**	**0.00376**	**0.02645**	**0.05404**	**0.9587**	**0.9287**
	Logsig	**4-2-1**	**0.00463**	**0.00165**	**0.01591**	**0.03432**	**0.9662**	**0.9327**
		4-3-1	0.00193	0.00486	0.03453	0.06342	0.9186	0.8812
		4-4-1	0.00424	0.00193	0.02559	0.03585	0.9424	0.9017
		4-5-1	0.00589	0.00427	0.02469	0.06358	0.9213	0.8915
		4-9-1	0.00818	0.00337	0.03465	0.06131	0.9481	0.9165
	Purelin	4-2-1	0.00397	0.00369	0.03373	0.05172	0.9147	0.8823
		4-3-1	0.00858	0.00342	0.02297	0.05495	0.9254	0.8937
		4-4-1	**0.00851**	**0.00325**	**0.02978**	**0.05374**	**0.9592**	**0.9218**
		4-5-1	0.00843	0.00345	0.02935	0.05856	0.9347	0.9028
		4-9-1	0.00852	0.00397	0.02954	0.05652	0.9443	0.9115

Similarly the LRNN results are discussed below for Loisingha station. For Tan-sig 4-2-1, 4-3-1, 4-4-1, 4-5-1, 4-9-1 architectures were taken into consideration for computation of performance revealed in Table 5. For Tan-sig function the best model architecture was found to be 4-3-1 which possessed MSE training value 0.00748, MSE testing value 0.00336, RMSE training value 0.02734, RMSE testing value 0.05796 and coefficient of determination value training 0.9972, testing value 0.9621. For Log-sig 4-2-1, 4-3-1, 4-4-1, 4-5-1, 4-9-1 architectures were taken into consideration for computation of performance. For Log-sig function the best model architecture was found to be 4-2-1 which possessed MSE training value 0.00133, MSE testing value 0.00285, RMSE training value 0.03511, RMSE testing value 0.05296 and coefficient of determination value training 0.9783 testing value 0.9588. For Purelin 4-2-1, 4-3-1, 4-4-1, 4-5-1, 4-9-1 architectures were taken into consideration for computation of performance. For Purelin function the best model architecture was found to be 4-4-1 which possessed MSE training 0.00975, MSE testing 0.00179, RMSE training 0.03122, RMSE testing 0.04448, and coefficient of determination training 0.9831, testing 0.9579.

Table 5
Results of LRNN at Loisingha for Scenario 2

Scenario 2	Transfer Function	Architecture	MSE		RMSE		R^2	
			Training	Testing	Training	Testing	Training	Testing
P_t, T_{max}, T_{min}, E_t	Tansig	4-2-1	0.00986	0.00125	0.13887	0.03598	0.9583	0.9267
		4-3-1	**0.00748**	**0.00336**	**0.02734**	**0.05796**	**0.9972**	**0.9621**
		4-4-1	0.00451	0.00371	0.02123	0.05891	0.9825	0.9566
		4-5-1	0.00489	0.00112	0.02211	0.03181	0.9813	0.9508
		4-9-1	0.00391	0.00955	0.01977	0.03090	0.9732	0.9495
	Logsig	**4-2-1**	**0.00133**	**0.00285**	**0.03511**	**0.05296**	**0.9783**	**0.9588**
		4-3-1	0.00937	0.00497	0.03061	0.07054	0.9375	0.9107
		4-4-1	0.00124	0.00654	0.03217	0.07908	0.9313	0.9002
		4-5-1	0.00513	0.00252	0.12459	0.05019	0.9724	0.9533
		4-9-1	0.00134	0.00653	0.03786	0.08080	0.9376	0.9016
	Purelin	4-2-1	0.00856	0.00367	0.02925	0.05802	0.9723	0.9415
		4-3-1	0.00991	0.00471	0.03148	0.06380	0.9601	0.9394
		4-4-1	**0.00975**	**0.00179**	**0.03122**	**0.04448**	**0.9831**	**0.9579**
		4-5-1	0.00972	0.00368	0.03117	0.06138	0.9528	0.9321
		4-9-1	0.00198	0.00512	0.03461	0.07079	0.9493	0.9277

Similarly Results of RBFN are presented in Table 6 in scenario various spread values were taken for simulation. Here spread values were considered within range of 0 to 1 i.e. preferably 0.2, 0.3, 0.5, 0.7 and 0.9 for predicating runoff from the considerable input parameters for mapping output. It was found that with a spread value 0.9 the RBFN showed best performance with architecture having 4-0.9-1 which possessed MSE training 0.00537 testing 0.00314, RMSE training 0.02317 testing 0.05589 and coefficient of determination training 0.9301 testing 0.9118

Table 6
Results of RBFN at Loisingha for Scenario 2

Scenario 2	Architecture	MSE		RMSE		R^2	
		Training	Testing	Training	Testing	Training	Testing
P_t, T_{max}, T_{min}, E_t	4-0.2-1	0.00582	0.00284	0.02816	0.02817	0.7826	0.7418
	4-0.3-1	0.00635	0.01175	0.02519	0.10851	0.7982	0.7799
	4-0.5-1	0.00822	0.00853	0.02867	0.09248	0.8448	0.7861
	4-0.7-1	0.00594	0.00497	0.02437	0.07049	0.9291	0.8827
	4-0.9-1	**0.00537**	**0.00314**	**0.02317**	**0.05589**	**0.9301**	**0.9118**

Results for Scenario 2 at Saintala using BPNN, LRNN and RBFN

For Tan-sig function in BPNN, 4-2-1, 4-3-1, 4-4-1, 4-5-1, 4-9-1 architectures were taken into consideration for computation of performance. For Log-sig function the best model architecture was found to be 4-2-1 which possessed MSE training value 0.00278, MSE testing value 0.00894, RMSE training value 0.05174 RMSE testing value 0.09270 and coefficient of determination for training 0.9607 and testing 0.9383. For Tan-sig and Purelin 4-2-1, 4-3-1, 4-4-1, 4-5-1, 4-9-1 architectures were taken into consideration for computation of performance. The detailed results are presented below in Table 7.

Table 7
Results of BPNN at Saintala for Scenario 2

Scenario 2	Transfer Function	Architec-ture	MSE		RMSE		R^2	
			Training	Testing	Training	Testing	Training	Testing
$P_{t,}$ $T_{max,}$, $T_{min,}$ E_t	Tansig	4-2-1	0.00283	0.00521	0.05461	0.04257	0.9491	0.9085
		4-3-1	0.00357	0.00217	0.05938	0.05283	0.9517	0.9133
		4-4-1	**0.00217**	**0.00171**	**0.04708**	**0.06458**	**0.9598**	**0.9208**
		4-5-1	0.00272	0.00901	0.05217	0.04347	0.9501	0.9116
		4-9-1	0.00267	0.00837	0.05154	0.04286	0.9499	0.9083
	Logsig	**4-2-1**	**0.00278**	**0.00894**	**0.05174**	**0.09270**	**0.9607**	**0.9383**
		4-3-1	0.00242	0.00116	0.04931	0.00136	0.9383	0.9185
		4-4-1	0.00309	0.00319	0.05566	0.05584	0.9505	0.9327
		4-5-1	0.00283	0.00655	0.05322	0.02559	0.9424	0.9234
		4-9-1	0.00269	0.00779	0.05107	0.02791	0.9358	0.9043
	Purelin	4-2-1	0.00297	0.00954	0.05932	0.03088	0.9537	0.9244
		4-3-1	0.00262	0.00347	0.05120	0.02167	0.9351	0.9026
		4-4-1	0.00269	0.00789	0.05117	0.02808	0.9273	0.8929
		4-5-1	**0.00283**	**0.01521**	**0.05461**	**0.02537**	**0.9595**	**0.9302**
		4-9-1	0.00357	0.02217	0.05938	0.01483	0.9317	0.8933

The LRNN results are discussed below for Saintala station. For Tan-sig, Log-sig, Purelin 4-2-1, 4-3-1, 4-4-1, 4-5-1, 4-9-1 architectures were taken into consideration for computation of performance. For Log-sig function the best model architecture was found to be 4-2-1 which possessed MSE training value 0.00632, MSE testing value 0.00612, RMSE training value 0.05130, RMSE testing value 0.07823 and coefficient of determination value training 0.9901, testing value 0.9827. Detailed results for other transfer functions are tabulated below in Table 8.

Similarly results of RBFN are presented in Table 9 with various spread values are taken for simulation. Here spread values were considered within range of 0 to1 i.e. preferably 0.2, 0.3, 0.5, 0.7 and 0.9 for predicating runoff from the considerable input parameters for mapping output. It was found that with a spread value 0.3 the RBFN showed best

Table 8
Results of LRNN at Saintala for Scenario 2

Scenario 2	Transfer Function	Archi-tecture	MSE		RMSE		R^2	
			Training	Testing	Training	Testing	Training	Testing
$P_{t,}$ T_{max},, $T_{min,}$ E_t	Tansig	4-2-1	0.00387	0.00173	0.06153	0.36293	0.9372	0.9085
		4-3-1	**0.00277**	**0.00529**	**0.05202**	**0.07634**	**0.9889**	**0.9772**
		4-4-1	0.00283	0.00998	0.05322	0.03159	0.9623	0.9454
		4-5-1	0.00293	0.00678	0.05406	0.09315	0.9797	0.9643
		4-9-1	0.00348	0.00341	0.05520	0.05950	0.9605	0.9528
	Logsig	**4-2-1**	**0.00632**	**0.00612**	**0.05130**	**0.07823**	**0.9901**	**0.9827**
		4-3-1	0.00313	0.00458	0.05594	0.06370	0.9647	0.9424
		4-4-1	0.00285	0.00879	0.05338	0.08293	0.9445	0.9385
		4-5-1	0.00418	0.01553	0.05846	0.10272	0.9734	0.9647
		4-9-1	0.00289	0.00567	0.05374	0.02381	0.9681	0.9523
	Purelin	**4-2-1**	**0.00238**	**0.00572**	**0.05136**	**0.02391**	**0.9893**	**0.9798**
		4-3-1	0.00312	0.00676	0.05587	0.02654	0.9577	0.9423
		4-4-1	0.00357	0.00541	0.05529	0.02325	0.9722	0.9622
		4-5-1	0.00182	0.00829	0.0426	0.02879	0.9636	0.9507
		4-9-1	0.00387	0.00173	0.06153	0.06293	0.9692	0.9585

performance with architecture having 4-0.3-1 which possessed MSE training 0.00214 testing 0.04255, RMSE training 0.05304 testing 0.03775 and coefficient of determination training 0.9429 testing 0.9125.

Table 9
Results of RBFN at Saintala for Scenario 2

Scenario 2	Architecture	MSE		RMSE		R^2	
		Training	Testing	Training	Testing	Training	Testing
$P_{t,}$ T_{max},, $T_{min,}$ E_t	4-0.2-1	0.04382	0.03821	0.04184	0.02339	0.9218	0.9063
	4-0.3-1	**0.00214**	**0.04255**	**0.05304**	**0.03775**	**0.9429**	**0.9125**
	4-0.5-1	0.00308	0.06891	0.05751	0.02567	0.9184	0.8953
	4-0.7-1	0.00343	0.04978	0.05867	0.02231	0.9027	0.8861
	4-0.9-1	0.00329	0.07897	0.06024	0.02454	0.8995	0.8606

From the above research we found that among two scenarios while evapotranspiration was added with rainfall, maximum and minimum temperature all rainfall-runoff correlation models performs best as compared to prediction without considering evapotranspiration.

Simulation

The graphs with best values for runoff from precipitation, maximum temperature, minimum temperature, evapotranspiration using BPNN with Log-sig transfer function at Loisingha shown in Figure 5. For Saintala watershed, the best value for runoff model is presented in Figure 6. The graphs below show how these best values results in the variation between the observed runoff and predicted runoff.

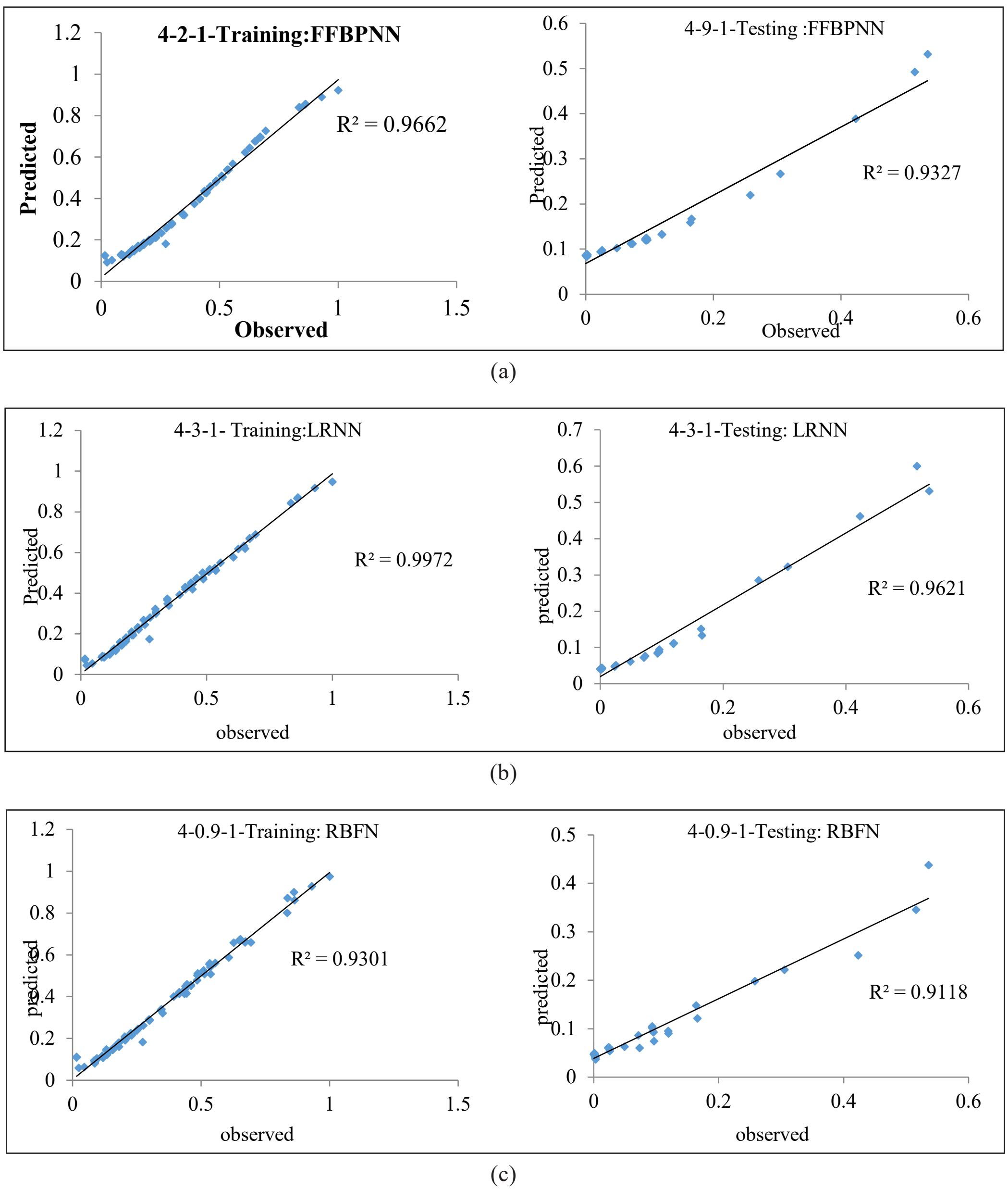
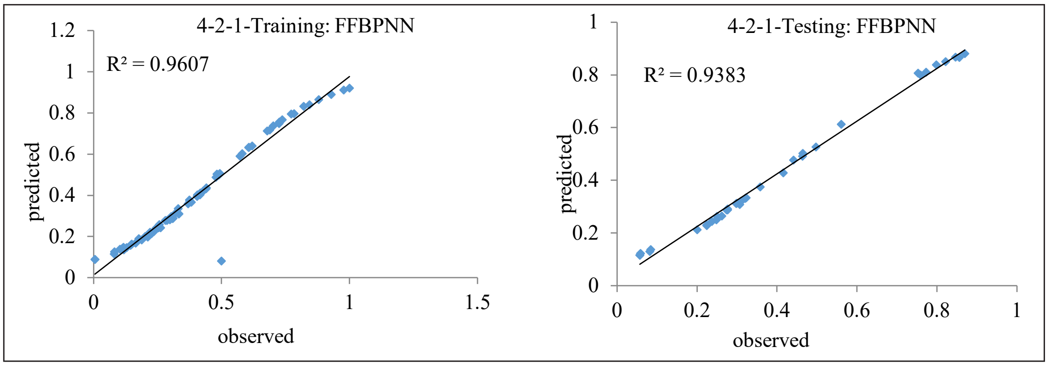
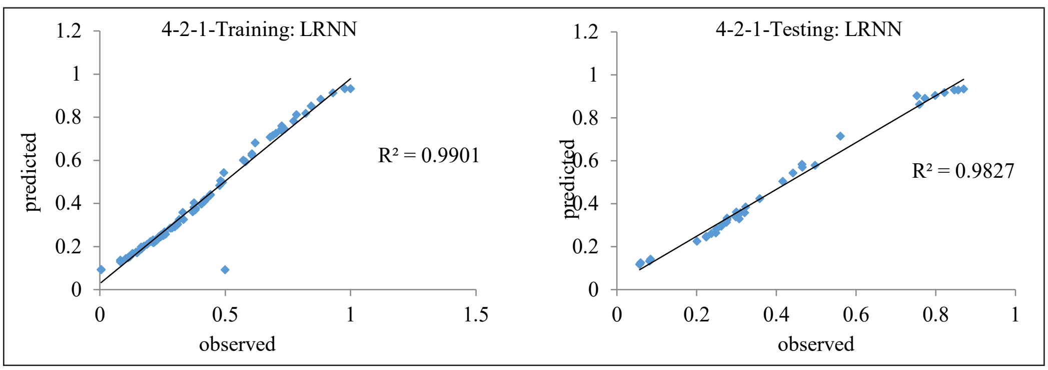


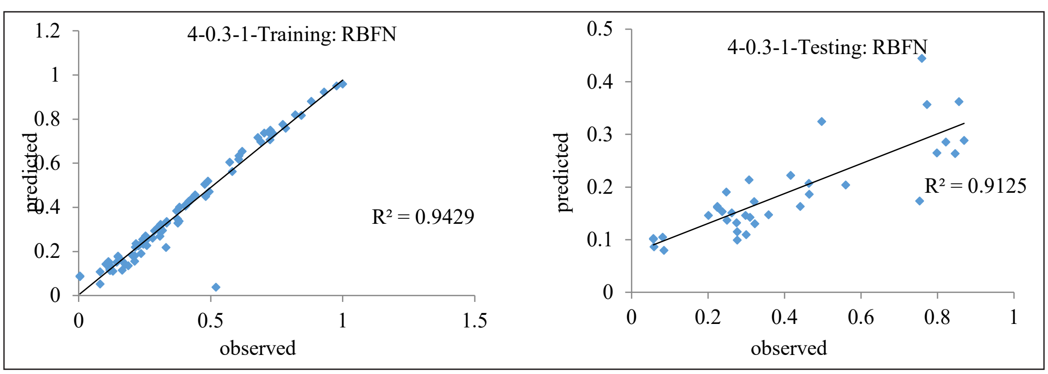
Figure 5. Best Observed V/s Predicted runoff model using BPNN, LRNN, and RBFN (Loisingha)



(a)



(b)



(c)

Figure 6. Best Observed V/s Predicted runoff model using BPNN, LRNN, and RBFN (Saintala)

Comparison of Model Performance for Two Watersheds

At Loisingha watershed among three neural networks with the evaluation criteria MSE, RMSE, and coefficient of determination LRNN performed best with architecture 4-9-1 following Tan-sig transfer function. At Saintala BPNN performed best among three networks with model architecture 4-3-1 using Purelin transfer function. The detailed result is tabulated below (Table 10). Since LRNN showed best efficiency the model would be utilized for predicting runoff of future uses nearest to the watershed and recommended for irrigation, and water resources department of Bolangir.

Table 10
Comparison of results for two watersheds

Scenario 2	Transfer Function	Tech-niques	Archi-tecture	MSE		RMSE		R^2	
				Training	Testing	Training	Testing	Training	Testing
P_t, T_{max}, T_{min}, E_t	Loisingha	BPNN	4-2-1 (Log-sig)	0.00463	0.00165	0.01591	0.03432	0.9662	0.9327
		LRNN	4-3-1 (Tan-sig)	0.00748	0.00336	0.02734	0.05796	0.9972	0.9621
		RBFN	4-0.9-1	0.00537	0.00314	0.02317	0.05589	0.9301	0.9118
	Saintala	BPNN	4-2-1 (Log-sig)	0.00278	0.00894	0.05174	0.09270	0.9607	0.9383
		LRNN	4-2-1 (Log-sig)	0.00632	0.00612	0.05130	0.07823	0.9901	0.9827
		RBFN	4-0.3-1	0.00214	0.04255	0.05304	0.03775	0.9429	0.9125

Assessment of Actual Runoff versus Simulated Runoff at Loisingha, and Saintala during Testing Phase

The variation of actual runoff vs. simulated or predicted runoff is revealed in Figure 7. Results shows that the estimated peak runoff were 306.2054 mm, 315.8574 mm, and 299.3439 mm for BPNN, LRNN, RBFN against the actual peak 328.3 mm for the watershed Loisingha, For watershed Saintala, the estimated peak runoffs were 344.5062 mm, 360.8081 mm, and 335.0335 mm for BPNN, LRNN, RBFN against the actual peak of 367.16 mm shown in Figure 8. This shows the significant potential of runoff and found to be useful for drought prone region with the predicted runoff index.

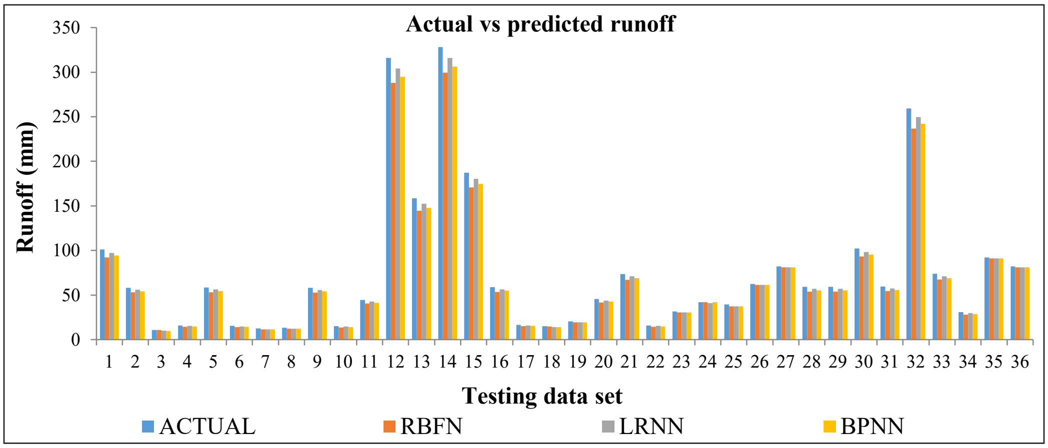


Figure 7. Actual v/s simulated runoff using BPNN, LRNN and RBFN at Loisingha in testing phase

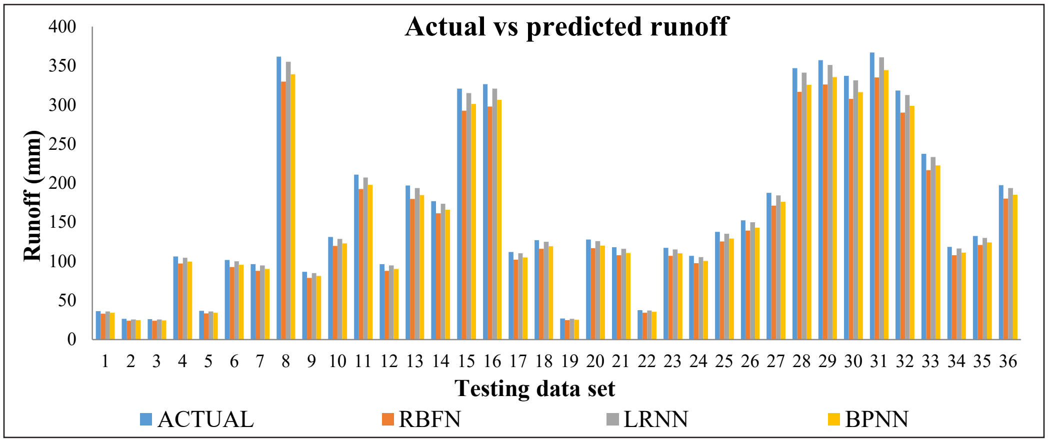


Figure 8. Actual v/s simulated runoff BPNN, LRNN and RBFN for Saintala in testing phase

CONCLUSIONS

Present study for predicting runoff is in fine agreement to model performance with evapotranspiration. Among two scenarios the prediction by evapotranspiration with all models performed best as compared to prediction without evapotranspiration. Overall performance with architecture 4-3-1 following Tan-sig transfer function, LRNN performed best. Model architecture 4-2-1 using Log-sig transfer function performed better with LRNN at Saintala. The best model performance was 0.9662 using BPNN, 0.9972 using LRNN and 0.9301 using RBFN at Loisingha. The influence of evapotranspiration was glowingly observed in watersheds for predicting runoff. Increase in evapotranspiration minimised the runoff potential for throughout the study area and *vice versa*. Since the area is within the region of scanty rainfall, the development of models will be helpful for assessing runoff. These results suggest the most suitable methods for developing environmental issues to estimate runoff in the two different watersheds of arid region. However, the combination technique needs to be investigated for improving integrated model techniques for future

ACKNOWLEDGEMENT

We thank the anonymous reviewers and editors for their useful suggestion to develop this manuscript.

REFERENCES

Asadia, S., Shahrabia, J., Abbaszadeh, P., & Tabanmehra, S. (2013). A new hybrid artificial neural networks for rainfall–runoff process modelling. *Neurocomputing, 121*, 470-480.

Chandwani, V., Vyas, S. K., Agrawal, V., & Sharma, G. (2015). Soft computing approach for rainfall-runoff modelling: A review. *Aquatic Procedia, 4*, 1054-1061.

Chen, J. Y., & Adams, B. J. (2006). Integration of artificial neural networks with conceptual models in rainfall-runoff modeling. *Journal of Hydrology, 318*(1-4), 232-249.

Dawson, C. W., & Wilby, R. L. (2001). Hydrological modelling using artificial neural networks. *Progress in physical Geography, 25*(1), 80-108.

Fahrmeir, L., Kneib, T., Lang, S., & Marx, B. (2013). *Regression models, methods and applications*. Heidelberg, Germany: Springer-Verlag.

Ghose, D. K., & Samantaray, S. (2018a). Modelling sediment concentration using back propagation neural network and regression coupled with genetic algorithm. *Procedia Computer Science, 125*, 85-92.

Ghose, D. K., & Samantaray, S. (2018b). Integrated sensor networking for estimating ground water potential in scanty rainfall region: challenges and evaluation. In B. B. Mishra, S. Dehuri, B. K. Panigrahi, A. K. Nayak, B. S. P. Mishra & H. Das (Eds.), *Computational intelligence in sensor networks* (pp. 335-352). Heidelberg, Germany: Springer.

Ghose, D. K., & Samantaray, S. (2018c). Sedimentation process and its assessment through integrated sensor Networks and Machine Learning Process. In B. B. Mishra, S. Dehuri, B. K. Panigrahi, A. K. Nayak, B. S. P. Mishra & H. Das (Eds.), *Computational intelligence in sensor networks* (pp. 473-488). Heidelberg, Germany: Springer.

Jain, A., & Srinivasulu, S. (2006). Integrated approach to model decomposed flow hydrograph using artificial neural network and conceptual techniques. *Journal of Hydrology, 317*(3-4), 291-306.

Javan, K., Lialestani, M. R. F. H., Ashouri, H., Moosavian, N. (2015). Assessment of the impacts of nonstationarity on watershed runoff using artificial neural networks: a case study in Ardebil, Iran. *Modelling Earth Systems and Environment, 1*(3), 22-31.

Kisi, O., Shiri, J., & Tombul, M. (2013). Modelling rainfall-runoff process using soft computing techniques. *Computers and Geosciences, 51*, 108-117.

Kothyari, U. C., & Singh, V. P. (1999). A multiple-input single-output model for flow forecasting. *Journal of Hydrology, 220*(1-2), 12-26.

Kumar, A. R. S., Sudheer, K. P., Jain, S. K., & Agarwal, P. K. (2005). Rainfall-runoff modelling using artificial neural networks: Comparison of network types. *Hydrological Processes: An International Journal, 19*(6), 1277-1291.

Lee, D. S., Jeon, C. O., Park, J. M., & Chang, K. S. (2002). Hybrid neural network modelling of a full-scale industrial wastewater treatment process. *Biotechnology and Bioengineering, 78*(6), 670-682.

Lungu, E. M. (1991). Stochastic characteristics of rainfall-runoff processes in southeast Botswana. *Hydrological Sciences Journal, 36*(5), 423-434.

Samantaray, S., & Ghose, D. K. (2018). Evaluation of suspended sediment concentration using descent neural networks. *Procedia Computer Science, 132*, 1824-1831.

Shamseldin, A. Y., O'Connor, K. M., & Nasr, A. E. (2007). A comparative study of three neural network forecast combination methods for simulated river flows of different rainfall-runoff models. *Hydrological Science Journal, 52*(5), 896-916.

Shoaib, M., Shamseldin, A. Y., & Melville, B. W. (2014). Comparative study of different wavelet based neural network models for rainfall–runoff modelling. *Journal of Hydrology, 515*, 47-58.

Smith, J., & Eli, R. B. (1995). Neural-network models of rainfall runoff process. *Journal of Water Resources Planning and Management, 121*(6), 499-508.

Srinivasulu, S., & Jain, A. (2006). A comparative analysis of training methods for artificial neural network rainfall-runoff models. *Applied Soft Computing, 6*(3), 295-306.

Tokar, A. S., & Johnson, P. A. (1999). Rainfall-runoff modelling using artificial neural networks. *Journal of Hydrologic Engineering, 4*(3), 232-239.

Zakermoshfegh, M., Ghodsian, M., Neishabouri, S. A. A. S., & Shakiba, M. (2008). River flow forecasting using neural networks and auto-calibrated NAM model with shuffled complex evolution. *Journal of Applied Science, 8*(8), 1487-1494.

Zhao, R. J. (1992). The Xinanjiang model applied in China. *Journal of Hydrology, 135*(1-4), 371-381.

Pertanika J. Sci. & Technol. 27 (4): 2265 - 2279 (2019)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Prediction of Groundwater Contaminants from Cattle Farm using Visual MODFLOW

Mohammad Nazri Ebrahim, Hasfalina Che Man*, Mohamed Azwan Mohamed Zawawi and Muhammad Hazwan Hamzah

Department of Biological and Agricultural Engineering, Faculty of Engineering, Universiti Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia

ABSTRACT

Livestock operation activities such as cleaning operation, feeding, milking and manure disposal are potential sources of contaminants into nearby surface and groundwater. In this study, the number of wastes generated from a cattle farm in Ladang 16 UPM, Serdang Selangor was estimated. Two monitoring wells were constructed at the site for groundwater quality monitoring assessment. The concentration of pollutants such as Potassium, Nitrate, and Copper was used in the simulation as an initial waste state. The simulation was conducted using Visual MODFLOW Software to predict the contaminants in groundwater. The aim was to predict the concentration of the pollutants distributed in groundwater and surface water sources in 365 days. Results of MODFLOW simulation showed that the flow of groundwater was in the direction towards the pond. The concentrations of Potassium, Nitrate, and Copper were predicted to accumulate in the groundwater to the pond within a year but the values were still below the drinking water standard. The groundwater contaminants could be due to seepage from the manure storage basin through subsoil into the shallow aquifer.

Keywords: Cattle farm, groundwater contaminants, monitoring well, visual MODFLOW

INTRODUCTION

The development in urbanization causes the increasing demand for water that has influenced groundwater to become one of the significant sources. Agriculture field development such as milk production, fertilizer disposal, livestock manure, and poultry production is currently growing rapidly. Nutrients enrichment such as

ARTICLE INFO

Article history:
Received: 07 February 2019
Accepted: 19 July 2019
Published: 21 October 2019

E-mail addresses:
emohammadnazri@gmail.com (Mohammad Nazri Ebrahim)
hasfalina@upm.edu.my (Hasfalina Che Man)
mohdazwan@upm.edu.my (Mohamed Azwan Mohamed Zawawi)
hazwanhamzah@upm.edu.my (Muhammad Hazwan Hamzah)
* Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

Nitrate, Copper, and Potassium. which are recognized as pollution, could result in an adverse effect on groundwater and surface water ecosystem (Adeoye et al., 2017). The main sources of nitrate pollution in livestock areas include livestock yard (that includes barnyards, holding areas, and feedlots), manure storage lagoons, and cropland receiving manure (Harter et al., 2002) areas which are the main sources of animal wastes within livestock facilities. Livestock waste contains nitrogen in both inorganic and organic compounds. The inorganic fraction is equivalent to the N emitted in urine and is usually greater than the organic one. Microbial action decomposes wastes containing organic nitrogen into ammonia, which is then converted into nitrite and nitrate. Nitrite is effectively oxidized to nitrate, so nitrate is prevalent in decayed wastes (Kumar, 2013). Nitrate can easily migrate through soil layers as the Nitrate-containing compound is generally soluble (Gollenhon & Caswell, 2000). Improper management of these compartments results in manure loss, which then leaches into the subsurface. The extent of leaching is dependent upon the climate, soil type, present nitrogen mass and hydraulic loading. Studies document nitrate leaching to groundwater under varying conditions, causing nitrate contamination of groundwater. Additionally, various studies from different regions in the United States and somewhere else, record manure effects to groundwater quality (Redding, 2016; Yamin et al., 2015). There are documented examples of manure impacts on groundwater quality in Washington State. In the Lower Yakima Valley, an U.S. Environmental Protection Agency (2011 and 2013a, 2013b) investigation concluded that dairy manure contributed to groundwater contamination of the local unconfined aquifer. There is currently no documented literature available on groundwater monitoring pollutants in Malaysia.

Based on the toxicity, the maximum contamination level (MCL) for drinking water has been set at 50 mg/L nitrate ion (NO_3) (equivalent to 11.3 mg/L nitrate-nitrogen, NO_3-N) by the World Health Organization (WHO), and 10 mg/L NO_3 – N by the U.S. Environmental Protection Agency (2003) and WHO (2004). Thus, nitrate is considered as an indicator parameter for assessing the extent of pollution in the vicinity of a facility. Other parameters that indicate pollution include ammonia, nitrite, pH, TDS, P, SO_4^{2-}, and alkalinity (Sahoo et al., 2016).Therefore, understanding of nitrate dynamics and other pollutants is important for managing and controlling potential groundwater pollution.

Groundwater monitoring provides a direct assessment of impacts to groundwater quality from land uses and is an important tool for determining how effective manure management practices are being implemented and thus minimizing impacts on groundwater. Groundwater monitoring is also an effective verification tool used to help evaluate the fate and transport of nitrate and other pollutants in the subsurface (Redding, 2016). Effective monitoring includes site selection criteria along with a selection of sampling points and parameters which can identify the source and extent of the contamination (Sahoo et al., 2016). Monitoring also provides an assessment of manure management practices.

Many groundwater models are available to assist in projecting impacts to groundwater quality. Some require intensive site-specific data, which typically generate more accurate results. The U.S. Geological Survey (USGS) developed MODFLOW in 1984 as a standard three-dimensional, finite difference groundwater model suitable for larger watershed assessments (Mc Donald & Harbaugh, 1984). It was originally developed in 1984, but has since been modified from being solely a groundwater flow model to now include contaminant transport, unsaturated zone transport, and water use by vegetation, as well as other capabilities. MODFLOW is a data-intensive model that requires accurate knowledge of environmental conditions. Visual MODFLOW (Waterloo Hydrogeologic, Kitchener, Ontario, Canada) is an application available to simulate groundwater flow in a wide range of natural systems. It is also able to determine the groundwater behaviour in connection with the river and lake as well as the estimation of contaminants transportation. The groundwater flow and contamination transport models are important in determining the sustainability of groundwater. For the purpose of modeling, data such as rainfall data, recharge rate, evaporation rate and parameters such as the concentration of contaminants are needed for estimating the transportation of contaminants by the groundwater.

Saghravani and Mustapha (2011) conducted a study on the movement of phosphorus leaching into groundwater using MODFLOW to predict subsurface and surface migration of pollution within 10 years. In addition, the application of MODFLOW to predict the pollution from poultry farm was conducted by Aderemi et al. (2014). The results of their study showed the simulation and prediction of phosphorus and nitrates concentration in the aquifer if indiscriminate and over application of poultry manure to Minna soil was continued. The simulation by MODFLOW was able to analyse the direction flow of the contaminants in groundwater. The MODFLOW can be used as a screening tool to indicate when groundwater is at risk of contamination. These tools provide an estimate of impacts on groundwater quality for the management of water resources and environmental protection.

The purpose of this study is to monitor the contaminants in groundwater at Ladang 16, UPM due to livestock operation activities. The sources of pollution are located on an elevated hillside and the spread of pollutions is expected to be more distributed as the groundwater carries the pollutant downhill (Saghravani, 2009). A river and lake were located less than 10 m from the source of pollution and thus needed to be monitored. Simulation using Visual MODFLOW is to predict the fate and transport of contaminants produced from the manure storage basin on groundwater quality on surface water and groundwater quality.

MATERIALS AND METHODS

General Description of Study Area

This study was conducted at Ladang 16, Universiti Putra Malaysia (UPM), Serdang, Selangor Darul Ehsan located at 2.991975° N, 101.733170° E. The 150,000 m^2 study area is located within elevation range between 52.5 m and 40 m above the Mean Sea Level (Figure 1). The population of dairy cattle and beef cattle in Ladang 16 was 43 and 135, respectively. The manure produced by a single cow was approximately 5-10 kg per day. The water used for cleaning purposes was estimated from 25 L to 30 L for each cattle per day. The wastewater from cleaning flowed directly into a retention pond. Two monitoring wells were constructed about 50 m to the cattle farm at A and next to a retention pond at location B (Figure 1). The monitoring wells MW1 and MW2 were located at 2.993113°N, 101.732°E and 2.99248°N, 101.7332°E.

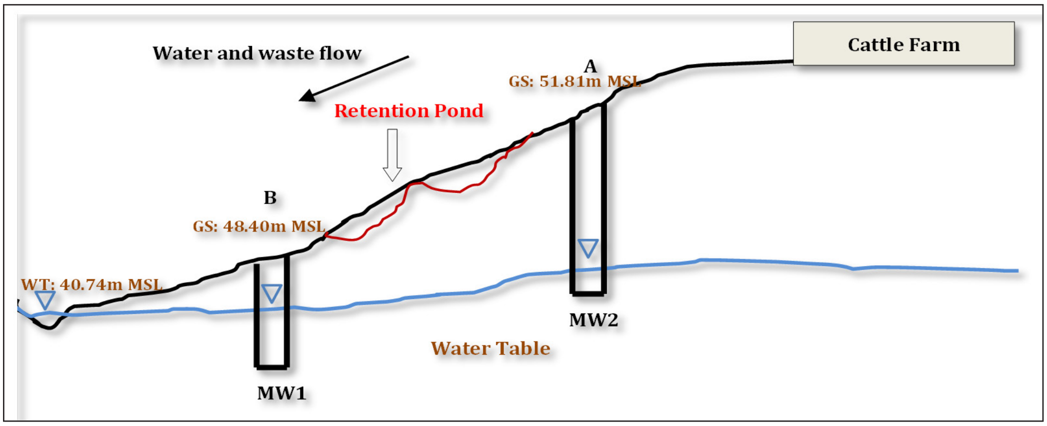


Figure 1. Cross-section view of the study area

Determination of Well Location

The well lithology and groundwater table were determined prior to the drilling of the borehole. Before that, a resistivity test was conducted to locate suitable areas for monitoring wells. The water samples were collected from MW1, MW2, pond, and river to be analysed in the laboratory for selected contaminants such as Copper, Potassium and Nitrate and *in-situ* groundwater analysis was measured at the monitoring well such as conductivity and pH by YSI meter.

Soil Resistivity Test

The soil resistivity was conducted to determine the geology type and possible groundwater storage for the area. This analysis was able to transversely map the soil layer based on the resistance values. A higher resistance was considered as hard rock and lower resistance was potentially to have more water. The Wenner array method was applied in this study. The experimental set-up of resistivity is shown in Figure 2. The cables were pulled to 100m to the left side and 100m to the right side from the centre at the Resistivity A and Resistivity B, respectively. The soil resistivity is calculated and recorded by ABEM Terrameter System. When all possible combinations of pairs of electrodes had been tested, the field data were processed using MODFLOW. Figure 2 shows the resistivity test at two different locations; location A is for MW2 and location B for MW1.

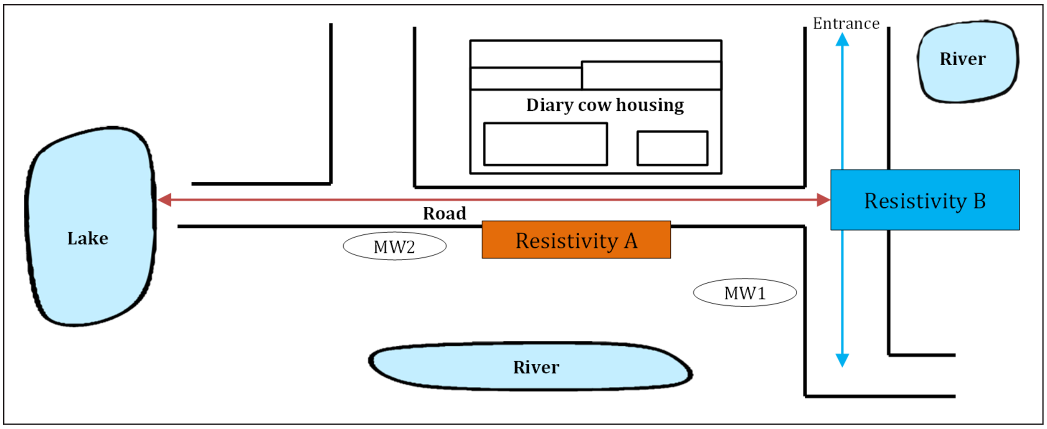


Figure 2. Resistivity test location A and B

Soil Sampling and Installation of PVC Pipe in the Borehole

Soil samples were collected at depths of 1.5 m, 3.0 m, 4.5 m, 6.0 m, 8.0 m and 10 m from each monitoring well during the borehole drilling. The 4 inches diameter of PVC pipes were installed as the main structure of the monitoring wells. The soil samples were classified by the Pipette method.

Simulation using Visual MODFLOW

The input data such as precipitation, evapotranspiration, river data and well lithology were collected through field measurement and from literature. The data required for MODFLOW simulation is summarized in Table 1.

Table 1
Data requirement in visual MODFLOW 4.2

Data	Value	References
Precipitation	2996.3 mm	(Saghravani et al., 2013)
Soil conductivity	Layer1(2.0 E-4 m/s) ,Layer2(3.2E-5 m/s), Layer 3(3.2E-10 m/s)	Environment base software.
Data of existing wells surrounding the site	Well Identification (Sel 1, Sel 49 and Sel 50)	Minerals and Geoscience Department Malaysia
Evapotranspiration	1498 mm	(Saghravani et al., 2013)

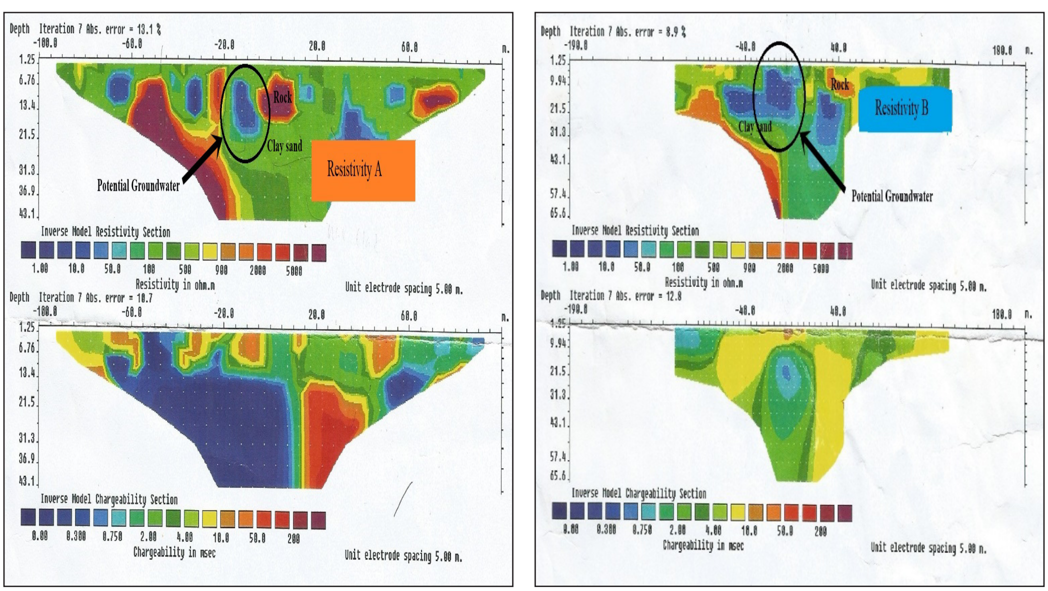
Water Quality Analyses

The portable sampler was used to collect groundwater at 10 m depth. The YSI portable meter was used to measure the pH, DO and turbidity. The groundwater and surface water (from lake and river) samples were collected from January 2017 to February 2017 for every two weeks. The water quality analysis was measured using DR/4000 Spectrophotometer within 190-1100 nm wavelength to measure the concentration of the nitrate, iron and sulphate in the water sample. Subsequently, water samples were also analysed using Inductively Coupled Plasma Optical Emission Spectrometry (ICP-OES) for heavy metals to obtain potassium and Copper concentration.

RESULTS AND DISCUSSION

Resistivity Test

As per results in Figure 3a, it was observed that the lower value of ohm-m indicates the presence of water at the depth of 10 m below the soil surface. A similar observation was reported by Meena (2011) that at -10 m from the central region of resistivity there were blue spots indicating groundwater with resistivity of 1 ohm-m. Figure 3b shows the availability of water at -40 m from the central region of resistivity. The resistivity value of topsoil was between 100-500 ohm-m (Meena, 2011). It shows that the type of composition is between clay sand and gravel mixture. The soil type provides a baseline for the selection of monitoring well, in addition to the water availability and the hardness of the soil type



(a) (b)

Figure 3. (a) Resistivity A to locate monitoring well 2 (MW2) and (b) Resistivity B to locate monitoring well 1(MW1)

beneath. Plus, the results for both resistivity A and resistivity B were compared with well log data from Pipette result along 10 m depth. The result is shown as the labelled structured.

Well Lithology from Pipette Method

The results from the Pipette method were calculated and the textural analysis based on the textural classification system. The result indicates that the topsoil consisted of 60% sand particles, 30% clay and 10% silt and the texture was sandy clay loam. The interception in the textural analysis was taken as the result of the soil classification. The result of soil classification was calculated using soil texture calculator. These results had been included as part of the inputs required by MODFLOW software such as the soil conductivity.

Water Quality Analysis

Based on Table 4, it shows the presence of contaminants of nitrate, copper and potassium in the river and pond that were suspected to be accumulated based on groundwater flow in Figure 1. Table 2 and Table 3 present the results of water quality in MW 1 and MW2. The results of the concentration of Potassium, Nitrate, and Copper in MW1 and MW2 were used in Visual MODFLOW to simulate the flow of transport in 365 days.

Table 2
Results of field analysis of groundwater quality

Parameter	Concentration in MW1	Concentration in MW2
Temp (°C)	28.4±0.85	27.8±1.25
DO (mg/L)	1.50±0.32	1.85±0.26
SPC (μs/cm)	189.73±11.9	50.7±3.25
TDS (mg/L)	123.18±6.22	32.83±2.22
Salinity (ppt)	0.088±0.005	0.02±0.0
pH	6.91±0.07	6.80±0.12

*values are means of the repeated experiments ± standard deviation

Table 3
Results of laboratory analysis of groundwater quality

Parameter	Concentration in MW1	Concentration in MW2
Turbidity (NTU)	*11.63±9.39	*7.10±2.02
Chemical (Inorganic)		
Ammonium (mg/L)	*5.80±1.33	*1.10±0.12
TSS (mg/L)	*207±113.76	*66.25±68.28
Nitrate (mg/L)	*1.43±1.85	*1.28±1.10
Sulphate (mg/L)	*15.03±4.94	*1.73±0.39
Iron (mg/L)	*2.78±1.56	*0.81±0.36
Potassium (mg/L)	7.09	5.26
Heavy Metals		
Copper (mg/L)	0.13	0.088
Lead (mg/L)	ND*	0.016
Microbiological		
Fecal Coliform (Cfu/100ml)	2	10

*ND= not detected; *values are means of the repeated experiments ± standard deviation

Table 4
Result of field Analysis of surface water quality

Parameter	Concentration in Pond	Concentration in River
Field Analysis		
Temp (°C)	*27.3±0.94	*27.1±1.37
SPC (µs/cm)	*88.48±12.60	*72.5±14.11
TDS (mg/L)	*57.36±8.13	*49.46±13.0
Salinity (ppt)	*0.04±0.01	*.03±0.01
Laboratory Analysis		
Turbidity (NTU)	*27.1±17.35	*13.45±14.29
Nitrate (mg/L)	1.7±1.95	1.4±1.74
Copper (mg/L)	0.091	0.075
Potassium (mg/L)	2.49	0.726
Fecal Coliform(Cfu/100ml)	>100	>50

The concentration of contaminants of Potassium, sulphate, and copper was high in MW1 as the source of pollution expected from the waste storage basin and the cattle farm. The concentration of contaminants in MW2 is much lower than MW1 presumably due to the infiltration effect of pollutants into the groundwater and downwards into the rivers and pond. Some of the waste could flow as a runoff from the higher elevated hillside to the river and pond (Figure 1). The contamination of potassium, nitrate and copper were measured inside both wells.

Figure 4 shows that the Potassium values were higher in MW1 and MW2 with 7.09 mg/L and 5.26 mg/L, respectively. The nitrate concentrations were at 3.28 mg/L in monitoring well 1 and 2.38 mg/L in monitoring well 2. The nitrate concentration was

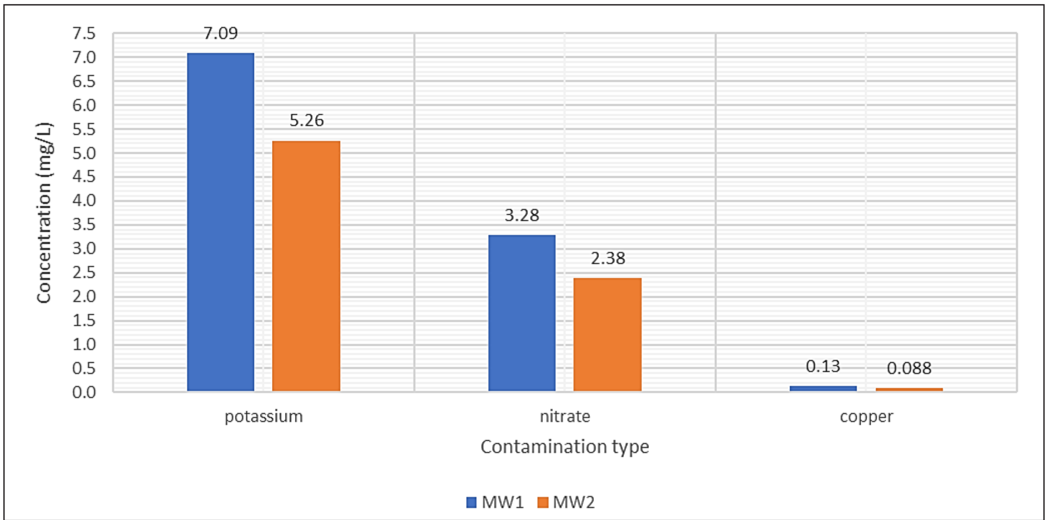


Figure 4. Concentration of contamination at MW1 and MW2

below the limits for drinking water. The contamination of copper was also low in MW1 which was 0.13 mg/L compared to 0.088 mg/ L inside MW2. The reduction of contaminant concentration from MW into pond and river indicates the lesser pollutant flows from higher elevated area to low. The results of contaminants, however, were higher in MW1 compared to MW 2.

Simulation using Visual MODFLOW

Generally, Visual MODFLOW predicts the flow and concentration of selected contaminants based on the sources release to the groundwater or aquifer during a simulation. Figure 5 shows the flow of groundwater was moving downwards in the direction of the pond. The result of the groundwater flow had proven to be in the same direction as shown in the cross-section view of the study area in Figure 1.

Validation is the process of evaluating the final product to check whether the software meets the expectations and requirements. In this study, validation was done by collecting water samples at MW1 and MW2 and analysing groundwater contaminants for Copper, Nitrate and Potassium after 1 year. Regression analysis was done to compare the observed concentration and predicted concentration from the model. The result was shown in Figure 9 for both monitoring wells MW1 and MW2. The correlation shows that R^2 is 0.8651 and 0.794 for MW1 and MW2, respectively. The coefficient of determination, R^2, is a measure of how well the regression model describes the observed data. From the R^2 value, this study shows that the predicted model was in a good fit with the observed data because R^2 shows the positive relationship and the nearer is R^2 to 1, the more accurate is the regression model (Schneider et al., 2010).

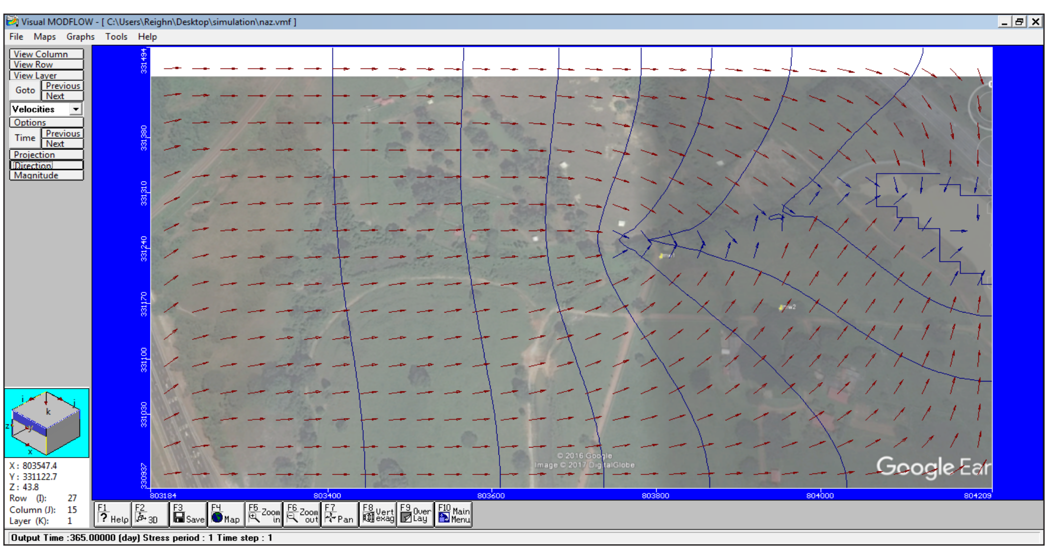


Figure 5. Groundwater flows in Ladang 16, UPM

Concentration of Copper

Figure 6 shows the predicted flow of Copper into the pond and the concentration was expected to be low in the area of cattle farm which was less than 0.0557 mg/L after 365 days due to the flow of groundwater that was distributed based on the high elevation to low elevation. The colour contour shows the flow and amount of copper within 365 days. As a result, copper was relatively low as it behaves in a different way from other heavy metals considered. Copper has been known to have good affinity with clay soil and can react with other chemicals in the soil to produce another compound (Aderemi et al., 2014). The distribution of concentration of Copper versus time is shown in Figure 6. The pattern of the graph shows that copper is predicted to decrease through time from its decrease initial concentration as it flows into the lake for 365 days. In conclusion, 0.0557 mg/L of copper is relatively low because the concentration for the toxicity of the groundwater affected is 20-50 mg/L.

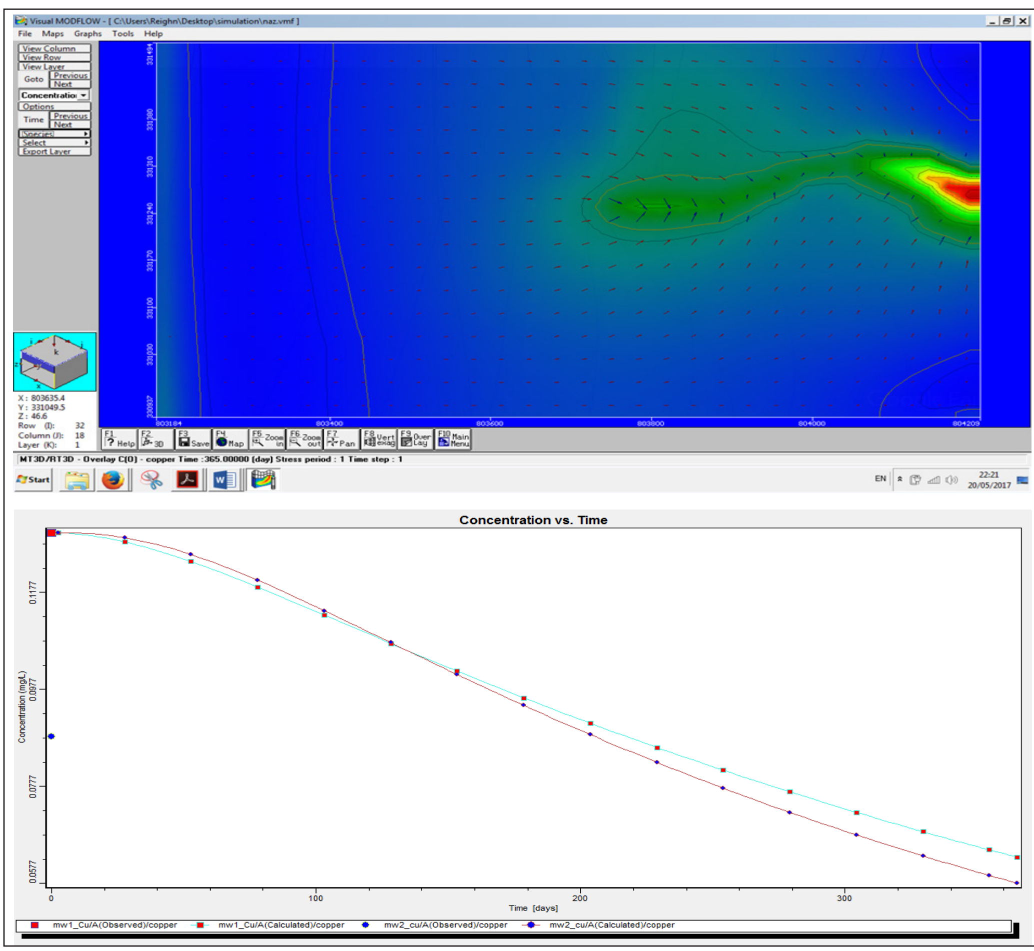


Figure 6. Copper Distribution in 365 days and its concentration by time

Concentration of Potassium

Potassium is available in manure and highly available in plants and can be used similar to K fertilizer application. Most potassium (> 70%) is excreted in urine, indicating the high solubility of the excreted K. Figure 7 shows the contour of the potassium distribution in 365 days, with the blue contour indicating the value of concentration was less than 3.15 mg/L. The potassium contamination was predicted to be at a high concentration level in the lake area. Within 365 days, the simulation showed the concentration of potassium was still below limits to give negative impacts to the environment.

Figure 7 shows the initial value of potassium was 7.15 mg/L and the concentration of potassium decreased with time. The lower concentration of potassium was predicted to be 3.15 mg/L and did not exceed the high level of standard concentration of potassium which was 5.2 mg/L. In conclusion, this groundwater can only be used for livestock operations such as drinking and cleaning because it will not harm the livestock yield, but it is not recommended to be used as drinking water because the concentration of potassium exceeds the standard concentration of 2.5 mg/L (WHO, 2009).

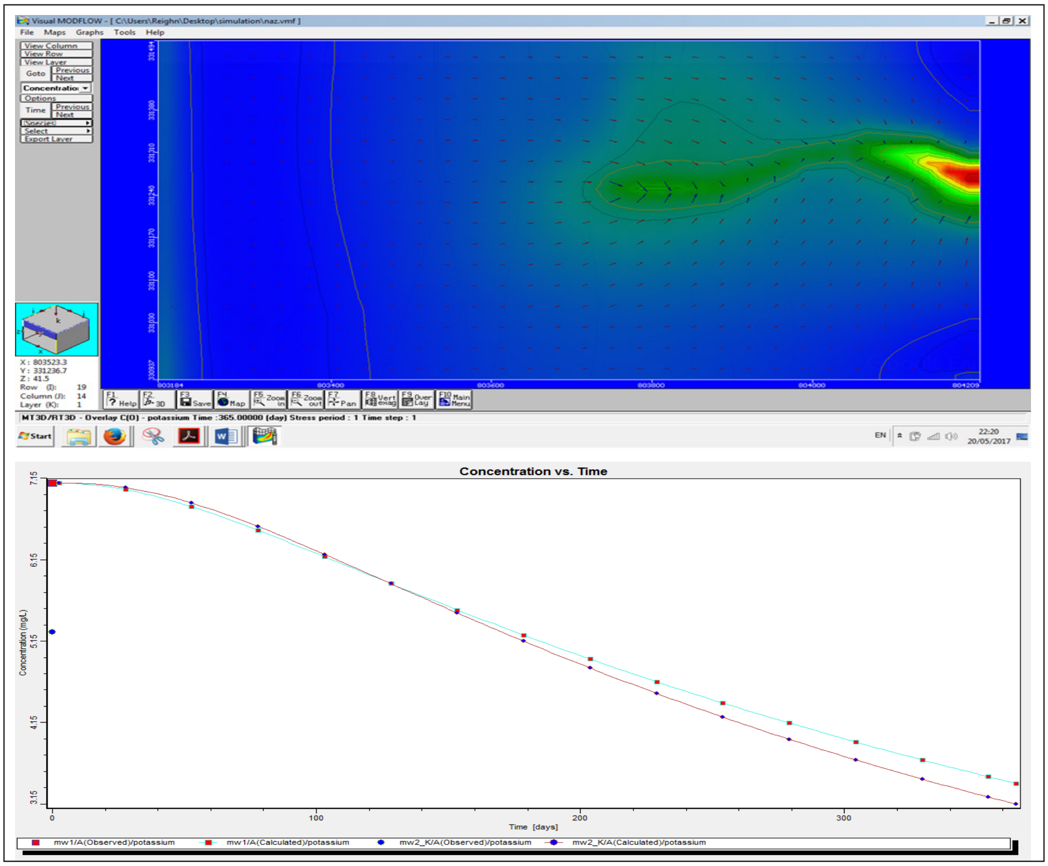


Figure 7. Potassium distribution in 365 days and its concentration by time

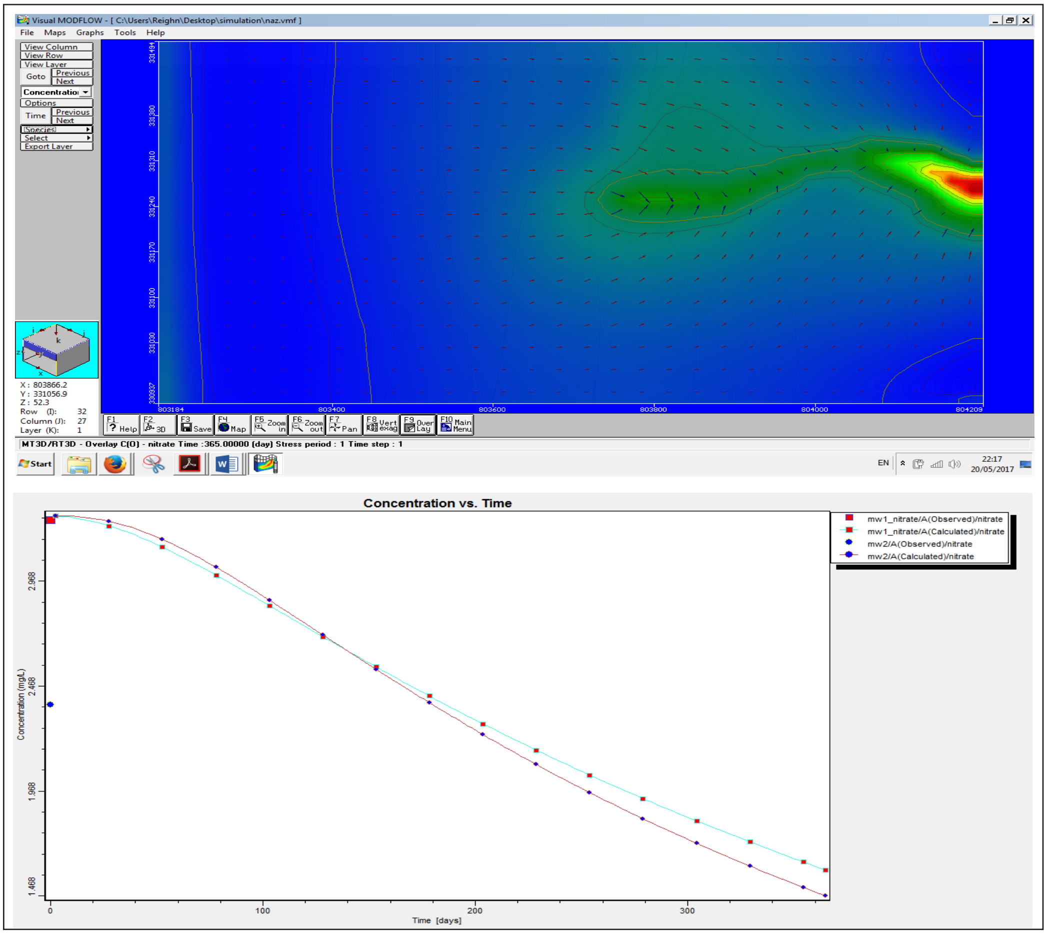


Figure 8. Nitrate distribution in 365 days and its concentration by time

Concentration of Nitrate

As shown in Figure 8, the nitrate concentration was flowing directly into the lake during the entire 365 days. The distribution of nitrate was predicted to be low in monitoring wells 1 and 2 in 365 days and the result of concentration showed that the nitrate decreased with time. Figure 9 shows the nitrate was 1.466 mg/L after 365 days compared to the initial concentration which is 3.28 mg/L. The result shows the pollution that was produced in the cattle farm had no adverse effect on the environment of groundwater because the concentration of nitrate did not exceed 4 mg/L.

In conclusion, the result of copper, potassium, and nitrate is relatively low in Ladang 16, UPM and the groundwater source can only be used for livestock operation and crop irrigation. The results indicate that within a year, the concentrations of nitrate and potassium in groundwater decreased by 55% and Copper decreased as well.

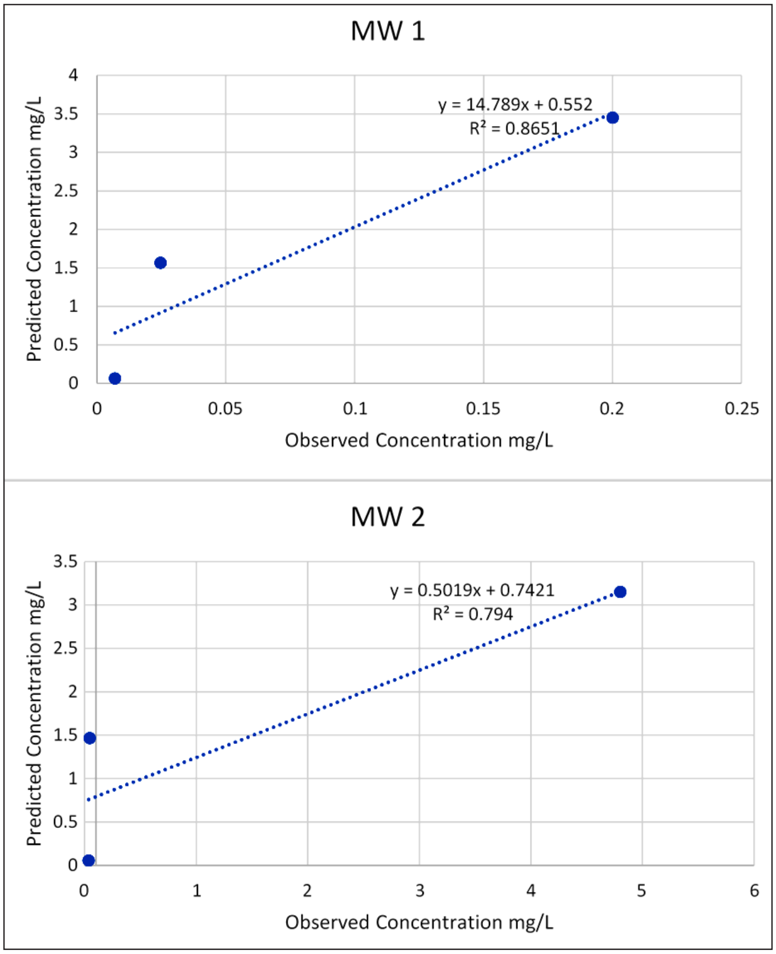


Figure 9. Predicted Concentration vs Observed Concentration

CONCLUSION

In this study, the quantity of waste generated from cattle farm has been estimated. Results from this study are useful for technical groundwater management to clearly identify suitable borehole locations for long term groundwater monitoring. In addition, this result can provide guidelines to determine suitable location for groundwater monitoring. Visual MODFLOW simulation is able to predict the pollutants fate in Ladang 16, UPM. It was observed that the concentrations of Copper, Potassium, and Nitrate were reduced within 365 days. Although the result showed that the pollution was low, actions in controlling and management of waste from any activities need to be done immediately. Furthermore, the result of simulation can be useful to assist the authorities in groundwater management as the main source of water supply in the future.

ACKNOWLEDGEMENT

The authors would like to acknowledge the financial support provided by TWAS (The World Academic Sciences) Research Grant (15-316 RG/ENG/AS_C/TWAS2015). Also thank to Mr. Tarmizi Ishak at University Agriculture Park for his assistance.

REFERENCES

Adeoye, P. A., Man, H. C., Soom, M. A. M., Thamer, A. M., & Oluwakunmi, A. C. (2017). Contaminants leaching from fresh poultry waste: a lysimeter study on sandy soils under tropical conditions. *Environment and Pollution*, *6*(1), 38-51.

Aderemi, P. A., Man, H. C., Soom, M. A. M., Mohammed, T. A., & Oluwakunmi, A. C. (2014). Groundwater quality of shallow wells on Nigerian poultry farms. *Polish Journal of Environmental Studies*, *23*(4), 1079-1089.

Gollehon, N., & Caswell, M. (2000). Confined animal production poses manure management problems. *Agricultural Outlook,* (274), 12-18.

Harter, T., Davis, H., Mathews, M. C., & Meyer, R. D. (2002). Shallow groundwater quality on dairy farms with irrigated forage crops. *Journal of Contaminant Hydrology, 55*(3-4), 287-315.

Kumar, L. (2013). A study of nitrate contamination in ground water of Delhi, India. Asian *Journal of Water, Environment and Pollution, 10*(2), 91-94.

McDonald, M. G., & Harbaugh, A. W. (1984). *A Modular Three-Dimensional Finite-difference Ground-Water Flow Model* (No. 83-875). Reston, VA: US Geological Survey,.

Meena, A. (2011). *Exploration of Ground Water using Electrical Resistivity Method* (Doctoral dissertation). National Institute of Technology Rourkela, Odisha, India.

Redding, M. (2016). *Manure and Groundwater Quality: Literature Review* (Publication No. 16-03-026). Washington, DC: Department of Ecology.

Saghravani, S. R. (2009). *Prediction of Phosphorus Concentration in an Unconfined Aquifer Using Visual Modflow* (Doctoral dissertation). Universiti Putra Malaysia, Selangor, Malaysia.

Saghravani, S. R., & Mustapha, S. (2011). Prediction of contamination migration in an unconfined aquifer with visual MODFLOW : a case study. *World Applied Sciences Journal, 14*(7), 1102-1106.

Saghravani, S. R., Yusoff, I., Mustapha, S., & Saghravani, F. S. (2013). Estimating groundwater recharge using empirical method : a case study in the tropical zone. *Sains Malaysiana, 42*(5), 553-560.

Sahoo, P. K., Kim, K., & Powell, M. A. (2016). Managing groundwater nitrate contamination from livestock farms: implication for nitrate management guidelines. *Current Pollution Reports, 2*(3), 178-187.

Schneider, A., Hommel, G., & Blettner, M. (2010). Linear regression analysis: part 14 of a series on evaluation of scientific publications. *Deutsches Ärzteblatt International, 107*(44), 776-782.

U.S. Environmental Protection Agency Science Advisory Board. (2011). *Reactive Nitrogen in the United States: An Analysis of Inputs, Flows, Consequences, and Management Options* (Report No. EPA-SAB-11-013). Washington, DC: EPA Science Advisory Board.

U.S. Environmental Protection Agency. (2003). *National Primary Drinking Water Standards*. Washington DC: Environmental Protection Agency.

U.S. Environmental Protection Agency. (2013a). *Relation between Nitrate in Water Wells and Potential Sources in the Lower Yakima Valley, Washington* (Report No. EPA-910-R-13-004). Washington, DC: Environmental Protection Agency.

U.S. Environmental Protection Agency. (2013b). *Monitoring Well and Installation and Data Summary Report, Lower Yakima Valley, Yakima County, Washington*. Washington, DC: Environmental Protection Agency.

World Health Organization. (2004). Recommendations: nitrate and nitrite. In *Guidelines for drinking-water quality* (3rd Ed.) (pp. 417-420). Geneva, Switzerland: WHO.

World Health Organization. (2009). *Potassium in Drinking-Water*. Geneva, Switzerland: WHO Document Production Services.

Yamin, M., Nasir, A., Sultan, M., Ismail, W. I. W., Shamshiri, R., & Akbar, F. N. (2015). Impact of sewage and industrial effluents on water quality in Faisalabad, Pakistan. *Advances in Environmental Biology, 9*(18), 53-59.

Pertanika J. Sci. & Technol. 27 (4): 2281 - 2304 (2019)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Spatial Pattern and Temporal Variability of Sea Level Anomaly and Geostrophic Current in the Eastern Indian Ocean from Satellite Altimetry

Julian Saputra[1,2]*, Jonson Lumban-Gaol[2], James Parlindungan Panjaitan[2] and Agus Saleh Atmadipoera[2]

[1]Graduate School of Marine Technology Program, Institut Pertanian Bogor, West Java 16128, Indonesia
[2]Department of Marine Science and Technology, Institut Pertanian Bogor, West Java 16128, Indonesia

ABSTRACT

The Eastern Indian Ocean has complex dynamic circulation systems affected by monsoonal wind circulation and climate variation. This research aimed to investigate the spatial and temporal variability of the sea level, and the geostrophic currents in the Eastern Indian Ocean using altimeter data. We used daily time series sea level anomaly data from 2004-2016 and applied time-series analysis of EOF. The highest four-modes were adopted. Both sea level anomaly and geostrophic zonal component had the sum of explained variance with 79.68% and 35.65%, respectively. The spatial pattern of the sea level anomaly showed dominant variability spread around the west coast of Sumatra and south coast of Java with positive and negative anomalies, as shown in first and second modes, while the third and the fourth mode did not show substantial spatial variability. The dominant temporal variation revealed semi-annual, annual and inter-annual periodicity. Furthermore, for the geostrophic zonal current, spatial pattern in the first to third modes showed high variability around the west coast of Sumatra and south coast of Java, while in the offshore region they showed positive-negative anomalies. In the fourth mode, the spatial pattern showed low variability. The dominant temporal variation revealed annual, semi-annual and inter-annual periodicity. Based on the spatial and temporal variation of the data, the variability in the study area might be associated with the dynamic of coastally trapped Kelvin wave, upwelling and an inter-annual anomaly of IOD and ENSO. It was also noticed that some patterns indicated as dynamic of SJC and SEC signals.

Keywords: Altimeter data, geostrophic zonal component, sea level anomaly, the Eastern Indian Ocean, time-series analysis EOF, variability

ARTICLE INFO

Article history:
Received: 16 January 2019
Accepted: 11 July 2019
Published: 21 October 2019

E-mail addresses:
julian.saputra75@gmail.com (Julian Saputra)
jonson_lumbangaol@yahoo.com (Jonson Lumban-Gaol)
jppanjaitan1@yahoo.com (James Parlindungan Panjaitan)
agus.itk@gmail.com (Agus Saleh Atmadipoera)
* Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

INTRODUCTION

Satellite altimetry has been developed comprehensively. The early ages on satellite altimetry history was the first launch of Skylab (1973-1974) that showed the capability of measuring geoid and oceanographic parameters. In 1978, the new generation of Seasat satellite was launched, making the improvement in quality of data, which was widely distributed to the scientists (Escudier et al., 2017). Long term altimeter missions were started when Topex/Poseidon (T/P) was launched in 1992 to the latest Jason 3 mission in 2016. They provided assimilation products to improve accuracy of altimeter data. The assimilation missions produced reliable spatial and temporal resolution data (AVISO, 2016).

The data describes sea level variation which is very useful for advanced analysis, such as to calculate the trend of sea level rise, to detect climate and weather, and to classify polar ice sheets (Fu & Cazenave, 2001; Mansawan et al., 2016; Rosmorduc et al., 2011). Understanding variability of sea level is very useful to predict potential fishing ground (Fitrianah et al., 2016; Lumban-Gaol et al., 2015; Syah et al., 2016). Sea level variability information can be used to detect gobal climate change e.g., Dipole Mode and El-Nino/La-Nina (Lim & Hendon, 2017; Syamsudin & Kaneko, 2013). Sea level data has been used widely to detect ocean dynamic, since it is directly related to ocean circulation such as geostrophic approximation (Fu & Cazenave, 2001; Le Traon, 2013). The understanding of geostrophic current is necessary, since the major currents of global ocean e.g., Gulf Stream, Kuroshio Current, and the Antartic Circumpolar Current are approximately included geostrophic current (Gill, 1982).

One of the data produced by satellite altimetry is Sea Level Anomaly (SLA) which plays a direct role in ocean circulation, especially geostrophic flow. Hence, geostrophic current variability can be derived from sea level topography (Joseph, 2014). Some research utilizing altimeter data had already been conducted. One of the first research using 3 months altimetric Seasat mission over the Antartic Circumpolar Current (AAC) was carried out by Fu and Chelton (1985). The result showed that the ACC circulation around the the Southern Ocean tended to be eastward acceleration. Another research was done by Susanto et al. (2001), predicted that sea level were related to the thermocline depth in the Indian Ocean. Sea level data from satellite altimetry had been extensively used to determine the dynamics and variability of the ocean, especially in a mesoscale region (Potemra & Lukas, 1999). One of the recent studies has been done by Syamsudin and Kaneko (2013). They used SLA to detect variability in the Eastern Indian Ocean (EIO) region. They found a lot of variability occurring in the EIO were influenced by phenomena such as El-Niño/La Niña, Indian Ocean Dipole (IOD) and regional monsoonal winds.

This research was conducted in the EIO, precisely at 100°E-120°E and 4°S-20°S (Figure 1). This area is fascinating because the variability that occurred are subjects to many influences e.g., wind stress anomaly, particularly the influence of monsoonal winds

(Schiller et al., 2009; Susanto et al., 2001; Wyrtki, 1961). Moreover, sea water flow in the EIO is also influenced by the Indonesian Throughflow (ITF) that passes through the Makassar Strait as its central passage (Gordon & Kamenkovich, 2010; Song et al., 2004; Susanto & Song, 2015). The annual cycle of ITF flow in the Makassar Strait is stronger in the southeast monsoon, July to September (Atmadipoera et al., 2016; Susanto et al., 2012). The circulation occurring in this region, besides ITF, is the South Java Current (SJC). It flows in the tropical Indian Ocean along Sumatra and Java coast, and move through the Savu Sea to reach Ombay Strait, which is one of the primary ITF pathways (Sprintall et al., 2009). Another circulation in the EIO is the South Equatorial Current (SEC). The SEC brings low salinity water which flowing westward, and recirculate in the East Gyral Current (EGC) circulating eastward (Bray et al., 1997; Feng & Wijffels, 2002; Meyers et al., 1995). During the semi-annual monsoon transition period (April-May and October-November), the westerly wind forcing generates Kelvin wave in the equator Indian Ocean and propagate eastward, forming Coastally Trapped Kelvin Wave (CTKW) along western Sumatra and southern Java coast (Clarke & Liu, 1993; Sprintall et al., 2000). Upwelling and downwelling process along these coasts may be affected by the waves (Susanto et al., 2001).

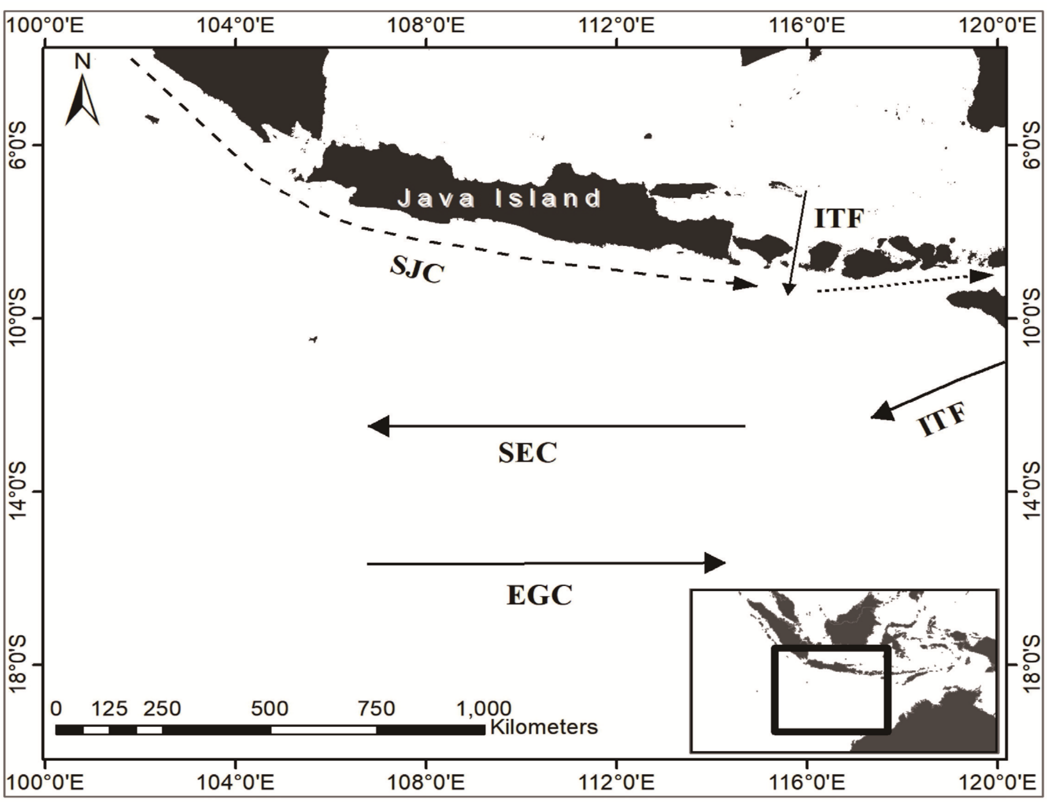


Figure 1. The surface current system in the EIO region. SEC: South Equatorial Current; SJC: South Java Current; EGC: Eastern Gyral Current; ITF: Indonesian Throughflow, adopted from Feng & Wijffels (2002) and Sprintall et al. (2010)

The interannual variability in the EIO region is prominently influenced by the IOD and El Niño-Southern Oscillation (ENSO) (Schott et al., 2009). The IOD is a pattern of internal variability in which sea surface temperature decreases in the eastern part of Indian Ocean, especially in the western coast of Sumatra, rising sea surface temperature in the western Indian Ocean region followed by high precipitation. The IOD occurs independently of the ENSO from the Pacific Ocean. Nevertheless, there is a link between IOD and ENSO. Usually, a negative IOD phase is followed by El Niño; likewise in the opposite phase, positive IOD is formed, followed by La Niña (Izumo et al., 2010). The ENSO plays essential roles in other oceanographic phenomena, i.e., sea level variation. During El Niño, the negative values of SLA occur near the south coast of Java, and positive values around the offshore region. The opposite situation occurs during La Niña, where the SLA values near the south coast of Java show positive values, while the offshore area has negative values (Syamsuddin et al., 2013).

Regarding the width of the coverage area, altimeter data is reliable for mesoscale studies, in which it has been constrained when utilizing in-situ data or model simulation. The limitation of using in-situ data is related to small area coverage, while model simulation results must be enhanced by validation (Fang et al., 2006). Considering the dynamic oceanographic processes and external influences from the Pacific Ocean and the effects of wind circulation in the Indian Ocean waters, the EIO region is very interesting for investigations. However, the collection of in-situ spatio-temporal data is challenging. Another challenge in using the model simulation is a validation procedure. Therefore, SLA data was utilized to produce variability in the EIO region. The purposes of this study were to investigate the spatial and temporal of SLA and geostrophic zonal current variability.

METHODS

The Data

The SLA data from satellite altimetry was used to meet the objectives of this study. The data was acquired in spatio-temporal form, better known as Delayed Time (DT) Mean Sea Surface Anomaly (MLSA). It is the integrated results of all altimeter missions such as HY-2A, Saral/AltiKa, Cryostat-2, Jason-1, Jason-2, T/P, Envisat, GFO, and ERS1/2. All the data from various missions were collected by *Segment Sol multimissions d'ALTimétrie, d'Orbitographie et de locali sation précise*/Data Unification and Altimeter Combination System (SSALTO/DUACS) (AVISO, 2016). Data could be downloaded on Copernicus Marine Environment Monitoring Service website (CMEMS http://marine.copernicus.eu/). The data spanning was 2004-2016 with daily temporal resolution. The SLA data was in gridded form with 0.25°x0.25° and was available in global coverage. The data coverage was 100°E-120°E and 4°S-20°S with the number of grids were 81x61, and 4749 layers representing time-series dimension.

Data preparation

The SLA data was available in global coverage. Cropping process was applied to achieve data in the EIO region only. The calculation in Stewart (2008) was followed to achieve geostrophic components. The surface geostrophic current estimation were calculated from the difference in water slope causing pressure differences. The zonal (u) and meridional (v) current components were obtained from the oceanic topographic gradient, gravitational force, and coriolis force. If it is assumed that density *(ρ)* and gravity (g) are constant at surface level, then geostrophic components can be calculated :

$$u = -\frac{g}{f}\frac{\partial \zeta}{\partial y};\ v = \frac{g}{f}\frac{\partial \zeta}{\partial x} \qquad (1)$$

where g is gravity, f is the coriolis force, and ζ is the sea level (Stewart 2008). The coriolis force equation can be written:

$$f = 2\Omega \sin\varphi \qquad (2)$$

where Ω is the angular velocity of the earth's rotation (rad/det), φ is the value of latitude.

The Coriolis force occurs due to the Earth's rotation, affecting the circulation into deflection to the right in the Northern Hemisphere and to the left in the Southern Hemisphere (Pond & Pickard, 1983; Tomczak & Godfrey, 2001). Although zonal and meridional current components were produced for geostrophic current, only zonal component was used in this study, because the dominant circulation in the EIO was west-east direction, as resulted from vector plot of geostrophic calculation derived from Conductivity, Temperature and Depth (CTD) data (Syamsudin & Kaneko, 2013).

Statistical process was applied to the data to determine the spatio-temporal mean and standard deviation of both on SLA data and geostrophic zonal component. The mean calculation was applied on SLA and geostrophic zonal component value of each grid, for all data both spatial and temporal. Meanwhile, the standard deviation was calculated by applying root-squared variance. After statistical processing, the climatological mean of SLA and geostrophic zonal component were calculated. The climatological mean was applied by averaging all data in the same month for all years.

Empirical Orthogonal Function (EOF) Analysis

Spatio-temporal variability is difficult to understand due to large coverage area and long time series data. Therefore, some techniques were applied to discover extreme variability from massive data. The idea was to compress the data into smaller values containing independent information. This can be done by applying EOF analysis. EOFs are a method for partitioning varian from a spatial data set that is distributed in time series, and they have been widely used in oceanographic data processing (Liu & Weisberg, 2007; Thomson

& Emery, 2014). EOF calculation was documented entirely and explained by Hannachi (2004) and Björnsson & Venegas (1999).

The principle of EOF analysis which is carried out in Hannachi et al. (2007) is to obtain the spatial (s) and temporal composition (t) denoted as:

$$X(t,s) = \sum_{k=1}^{M} c_k(t)\, \mathrm{u}_k(s) \tag{3}$$

where M is the number of modes resulting from the decomposition of spatial signals $\mathrm{u}_k(s)$ and expansion coefficients or temporal principal components $c_k(t)$. If $X(t,s)$ is applied to SLA data and geostrophic component data in gridded form, at time t and position s, then the discrete value t_i and grid position s_j can be written x_{ij} for $i = 1, \ldots, n$ and $j = 1, \ldots, p$. If written in the form of a matrix:

$$X = (x_1 x_2 x_3)^T = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix} \tag{4}$$

where $x_t = (x_{t1}, x_{t2}, \ldots x_{tp})^T$ describes the space in time t.

The limitation of this analysis regarding to variability patterns given by this EOF depict standing oscillation. Some limitation of EOF is the difficulty of physical interpretation. To overcome this shortcoming, the EOF method was modified to be Rotated EOF (REOF). This technique is also known as factor analysis as a factor rotation and aims to obtain a simple structure (Björnsson & Venegas, 1999).

EOF analysis produces variations in spatial patterns with eigenvector, describing the spatial pattern at all time range, the percentage of explained variance with eigenvalue, describing the contribution of variability in certain mode, and the temporal variation of Principal Components (PC) explaining how spatial pattern varies with time. EOF processing was carried out on SLA and geostrophic zonal component data throughout 2004-2016 to determine spatial and temporal variability in that period. EOF analysis on climatological mean data was also applied to determine spatial and temporal variability in the annual cycle. To determine how many modes would be used describing the variability, Jolliffe (2002) explained that the threshold for taking the cumulative explained variance was about 80%. So the certain numbers of modes that had 80% of cumulative explained variance were considered to explain the major variability in the EIO.

Power Spectral Density (PSD) analysis

Time series analysis was applied to determine dominant temporal variability from EOF analysis. Thus, EOF temporal variation results were analyzed by looking at the spectrum of energy density, where energy is defined as units per time domains (Thomson & Emery, 2014). Spectral energy analysis is better known as PSD analysis. PSD analysis was carried out to determine the energy value of each current signal per frequency. From this analysis, the relatively high power peak was shown to describe a dominant variability fluctuation in a certain frequency. Blackman and Tuckey developed a popular method to calculate PSD (as cited in Trauth, 2014, p. 158). In this study, PSD was calculated by determining the Fourier component of PC from SLA and geostrophic zonal component data generated from EOF in daily time scale. It was done to find more details from complex temporal variability occurring in the study area. The PSD analysis was calculated by:

$$PSD = \frac{X^*_{xx}(f)X_{xx}(f)}{f_s} = \frac{|X_{xx}(f)|^2}{f_s} \tag{5}$$

where $X^*_{xx}(f)$ is a complex conjugation of Fourier transforms from the autocorrelation function $X_{xx}(f)$ and f_s is the sampling frequency. The significant frequencies in PSD analysis were known by applying 95% confidence interval. The equation was written as:

$$\frac{v\tilde{X}_{xx}(f)}{X^2_{1-\alpha/2.v}} < X_{xx}(f) < \frac{v\tilde{X}_{xx}(f)}{X^2_{\alpha/2.v}} \tag{6}$$

where $\tilde{X}_{xx}(f)$ is the raw estimate of the observed time series, $X_{xx}(f)$ is the true spectrum, $X^2_{1-\alpha/2.v}$ and $X^2_{\alpha/2.v}$ are chi-square variables with $(1-\alpha)100\%$ confidence (Thomson & Emery, 2014).

RESULTS AND DISCUSSION

Mean and Standard Deviation

Mean and the standard deviation were calculated in both SLA and geostrophic flow data (Figure 2). The mean calculation of spatio-temporal data in Figure 2a showed that the maximum SLA value was 0.08 m with red color indicator, while the lowest value was 0.038 m with dark blue color indicator. The highest value was located around the coast of the Java Island, while the lowest value spread along the southern region far from the coast, around 100°E-106°E and 14°S-18°S. The overlayed geostrophic flow showed some eddies. There were two cyclonic eddies discovered at lower sea level, one in the area with coordinates 105°E-108°E and 15°S-16°S, and another area with coordinate 113°E-115°E and 16°S-17°S. The high SLA value with some vortexs showed around south coast of Java i.e., the Sunda Strait and another area around 108°E-110°E and 7°S-8°S.

The vortex occured in the Sunda Strait was probably because of the sea water circulation from the Java Sea and the Indian Ocean off west Java. In the Java Sea, the water circulation flowed westward, then turn southwestward through the Sunda Strait. On the other side, sea currents from the Indian Ocean around the west coast of Sumatra flowed southeastward along the southern coast of Java. Due to the convergence of the current in the Sunda Strait, it caused a vortex that occurred anticyclonically. Rahmawitri et al. (2016) said that sea water from the Java Sea flows to the Indian Ocean through the Sunda Strait in upper 5 m. The next vortex in 108°E-110°E and 7°S-8°S was downwelling phenomenon that happened caused of Kelvin wave and northwest monsoon.

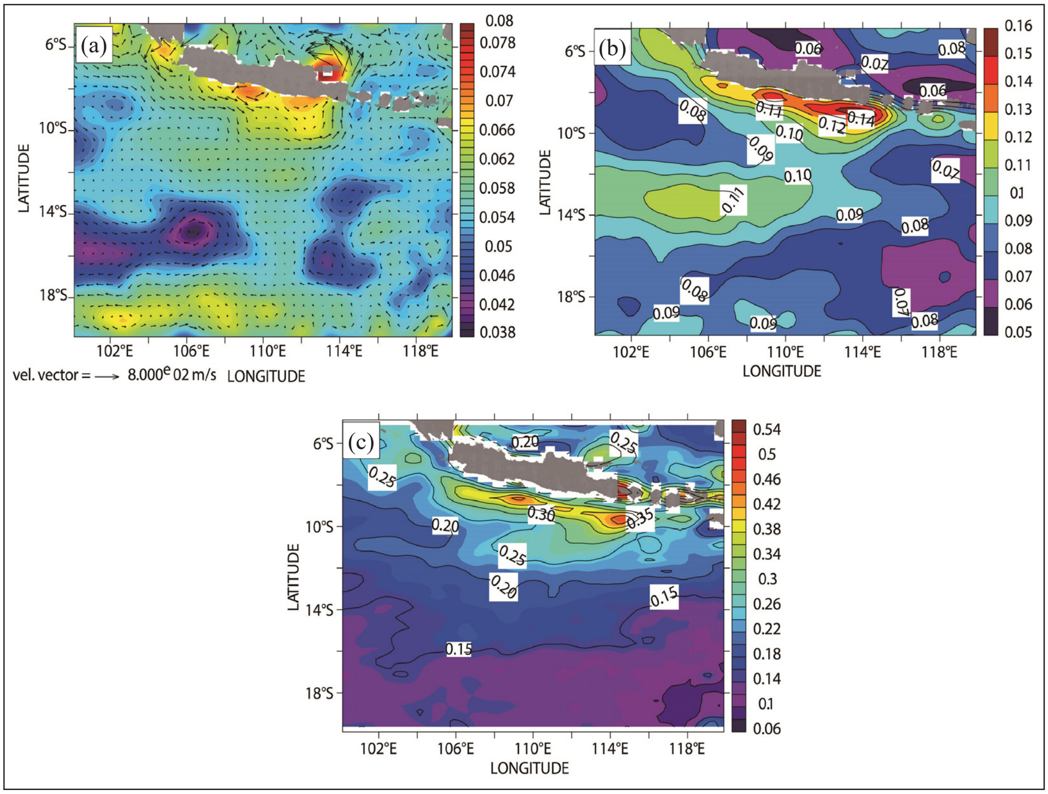


Figure 2. Mean and standard deviation of SLA and geostrophic current, (a) Mean of SLA and geostrophic current, (b) the standard deviation of SLA, (c) the standard deviation of geostrophic zonal component

The standard deviation of SLA data was presented in Figure 2b. In general, the higher standard deviation values were observed along west Sumatra and south Java coast, and around 10°S-15°S. However, in particular, the highest value was seen in the eastern part of south Java coast, precisely at 111°E-117°E and 8°S-11°S. It means that the highest difference of range from the average values were seen in that area. Meanwhile, the low standard deviations occurred at 114.5°E-120°E and 16°S-20°S. Furthermore, The geostrophic zonal component's standard deviations showed different aspects (Figure 2c). The higher standard deviations occurred along south Java coast, but the location was a little bit far from the coast. While the minimum values were mostly occurred at 14°S-18°S. The high number of standard deviations both in SLA and geostrophic zonal component near south coast of Java might be indicated as a high variability that occurred in that area.

Annual Cycle of SLA and Geostrophic Currents

The annual cycle of the SLA and geostrophic zonal component data were determined by applying EOF analysis of climatological mean data. This idea was generated from Fang et al., (2006). Spatial and temporal patterns were obtained with the two leading modes (Figure 3). The first mode accounted for 65.5% of the explained variance and the second mode accounted for 20.37% of the explained variance. The cumulative percentage of explained variance was 85.87%. In the first mode (Figure 3a), spatial patterns showed positive values in the lower part and negative values in the upper part of the study area. The centralization was seen near the south coast of Java, in the eastern part near Bali with the negative values. The different spatial patterns were shown in the second mode (Figure 3b), where there was a centralized spatial pattern with negative values at 111°E-118°E and 12°S-13°S. The PC graph of the first mode, presented in Figure 3c, described the temporal pattern showing the highest values occurring in August. Meanwhile, the PC graph of the second mode (Figure 3d) showed two positive peaks occurring in February and October, and one negative peak in May-June.

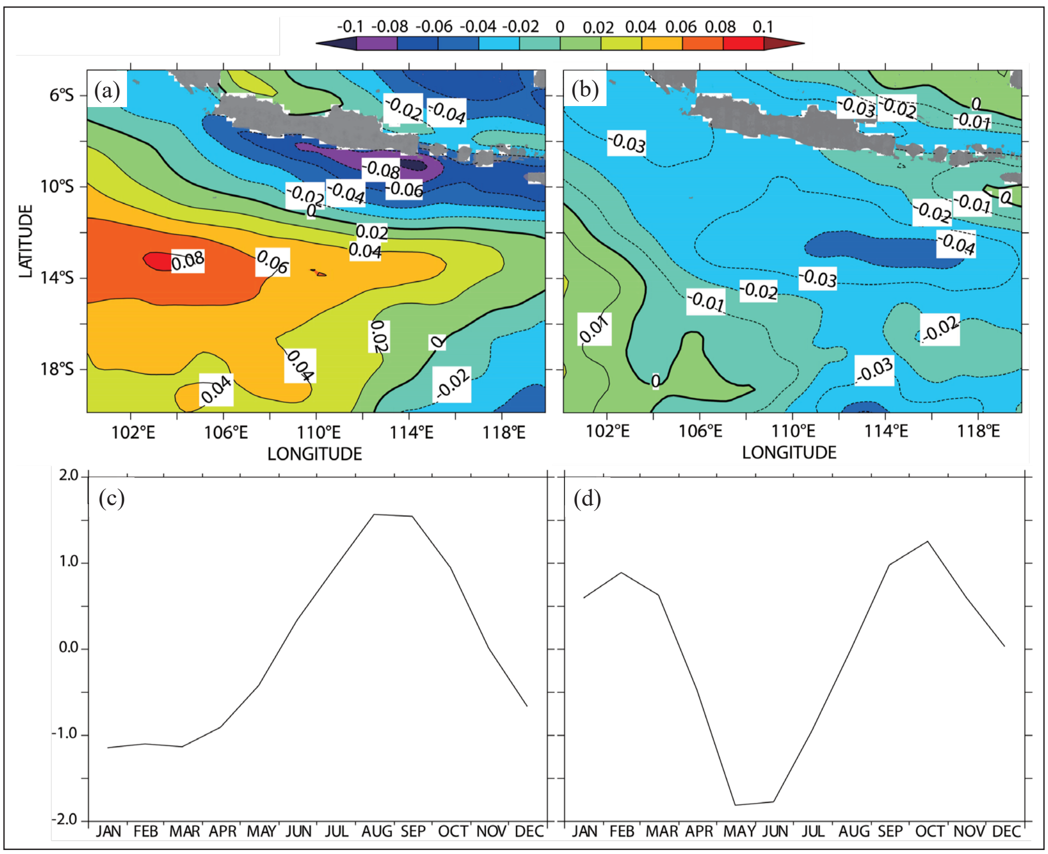


Figure 3. EOF analysis from SLA's climatological mean on two major modes, (a) spatial pattern the first mode, (b) spatial pattern the second mode, (c) PC graph of the first mode, (d) PC graph of the second mode

Annual cycle analysis of geostrophic zonal component was shown in the results of EOF processing on climatological mean. The results were displayed in two leading modes (Figure 4). The first mode accounted for 71.89% of the explained variance. The spatial pattern in the first mode (Figure 4a) showed negative values around the southern Java coast until 12°S, while the positive values distributed from 12°S to 20°S. Furthermore, in the second mode (Figure 4b) accounted for 9.74% of the explained variance, showing little variation with the positive values around the southern Java coast. Subsequently, in the temporal variation from the PC graph of the first mode (Figure 4c), the highest variability was seen in August and two negative peaks in January and December. Meanwhile, the temporal variation of the second mode from PC graph (Figure 4d) showed two positive peaks in May and November and two negative peaks in February and August.

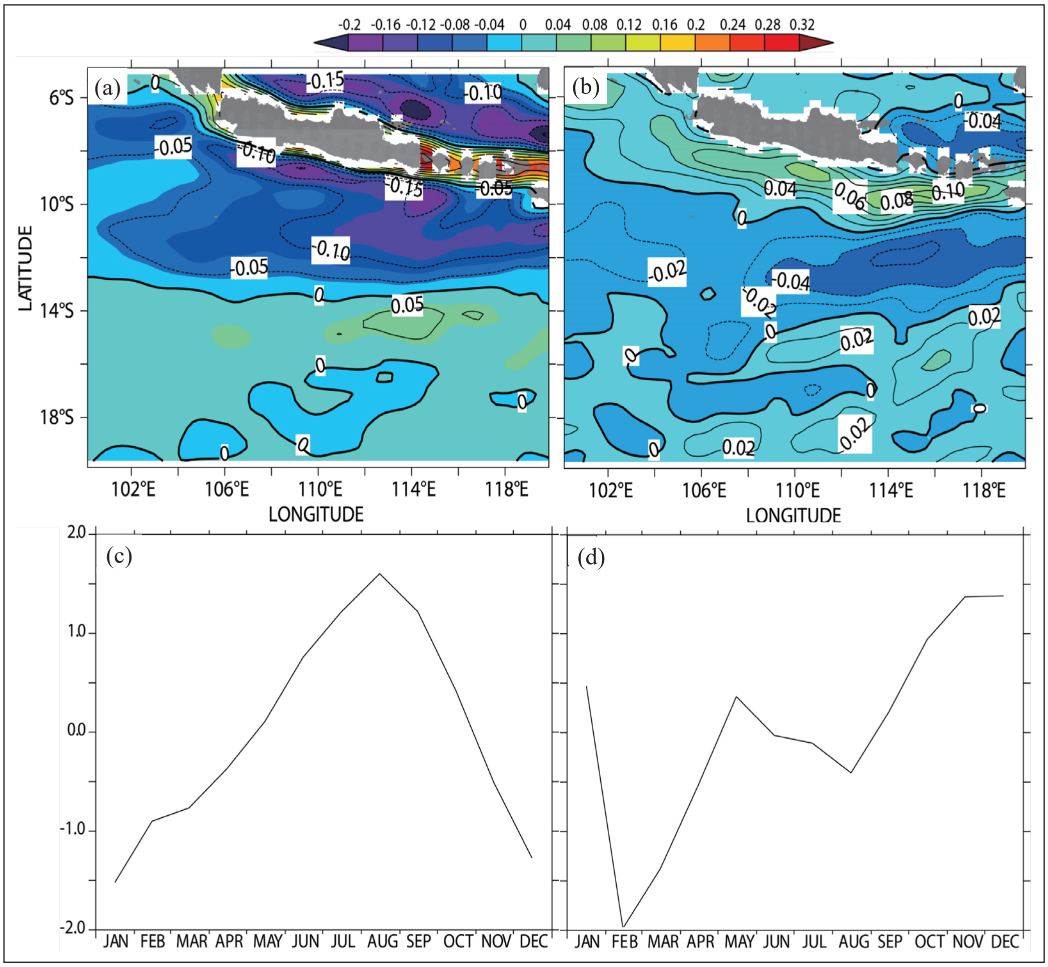


Figure 4. EOF analysis from the geostrophic zonal component's climatological mean on two major modes, (a) spatial pattern of the first mode, (b) spatial pattern the second mode, (c) PC graph of the first mode, (d) PC graph of the second mode

Variability of SLA and Geostrophic Zonal Component

The spatial and temporal SLA patterns were interpreted from the results of the EOF analysis in the EIO, displayed in the 4 biggest modes. The first mode (Figure 5) had the most substantial explained variance (52.31%). The spatial pattern showed positive values spread in the entire area. The highest values were seen along the west Sumatra and south Java coastz. The low eigenvector values formed a separate contour at remote locations from Java, which were pointed at 100°E-107°E and 10.5°S-16.5°S. In general, it was observed that the spatial pattern with high eigenvector values were located in the eastern part, while the low contour values were located in the western part of the study area (Figure 5a). The time series variation from PC was observed consistent oscillation's peak mostly happened in April, May, and November (Figure 5b). Furthermore, the energy fluctuations seen from PSD temporal analysis were revealed in the semi-annual (182 days), annual (356 days) and inter-annual (1024 days) periodicities (Figure 5c).

The different spatial pattern was shown in the second mode (Figure 6), accounting for 14.98% of the explained variance. The spatial pattern showed different phase oscillated in negative and positive values. The spatial pattern in this mode showed negative values along west Sumatra and south Java coast. A centered negative values located near the south of eastern Java were observed. The positive values spread in the southern part of the study

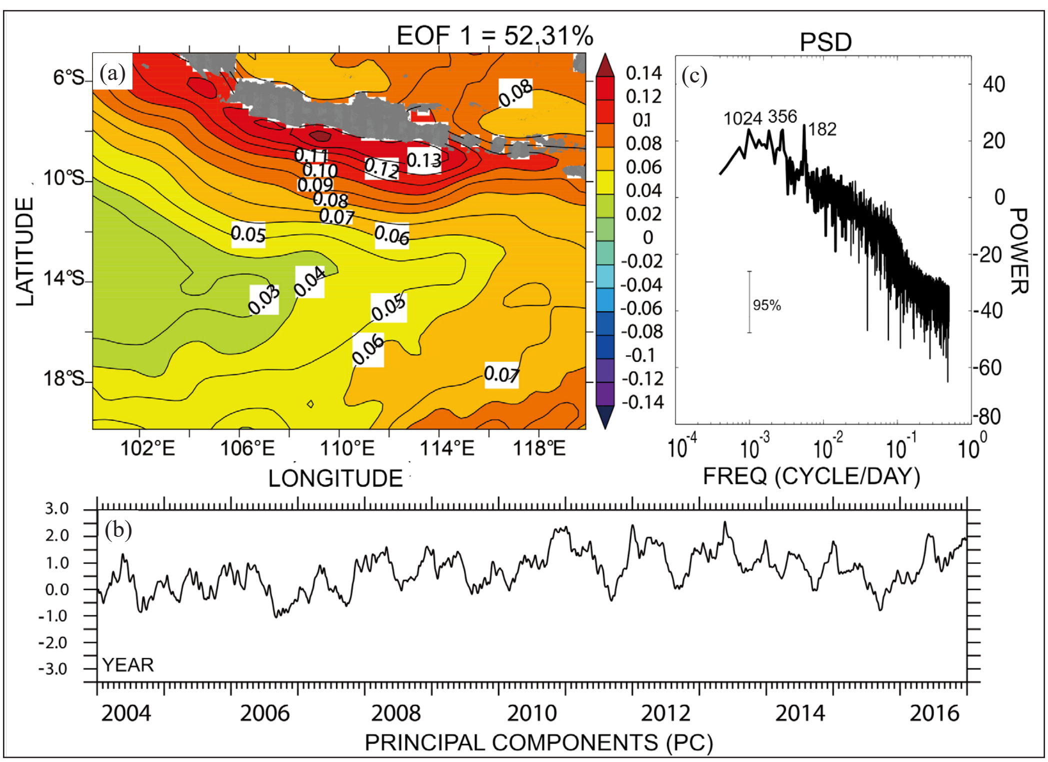


Figure 5. EOF analysis from SLA in the first mode (a) spatial pattern (b) PC graph (c) PSD analysis with 95% of confidence interval

area, especially in the offshore area (Figure 6a). Based on the spatial variation from PC and PSD. Clearly, the fluctuation happened with annual (356 days), inter-annual (1024 days) and semi-annual (186) periodicities (Figure 6b, 6c).

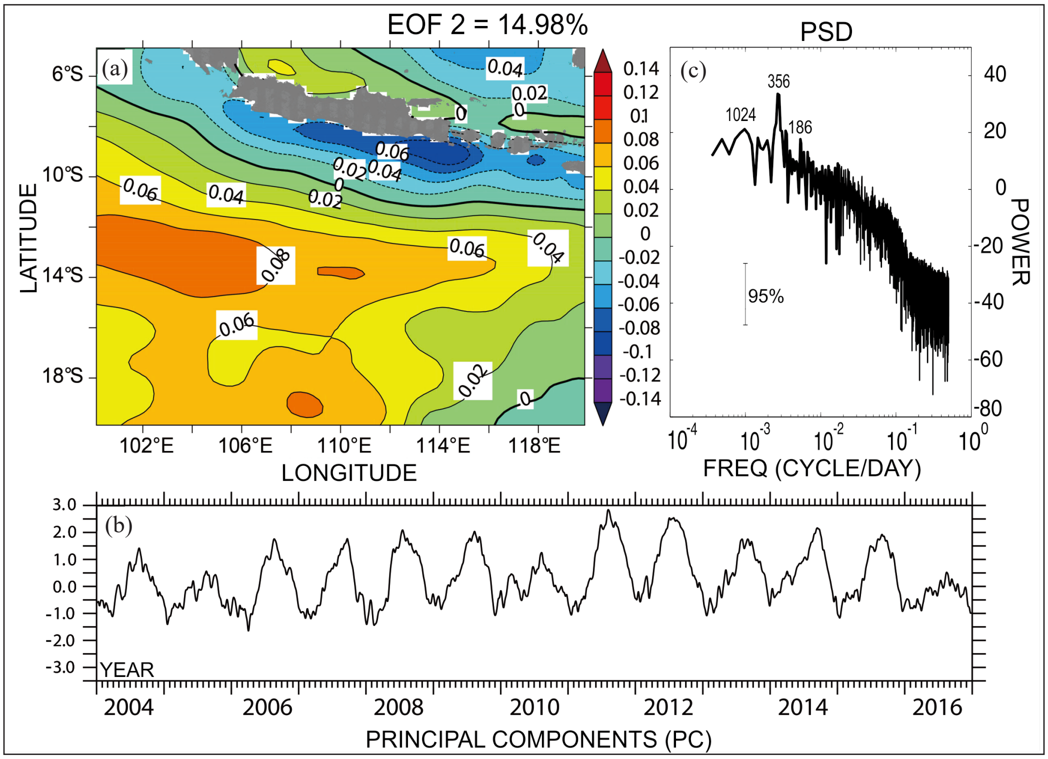


Figure 6. EOF analysis from SLA in the second mode (a) spatial pattern (b) PC graph (c) PSD analysis with 95% of confidence interval

The third and fourth mode were the minor modes, accounting for 7.84% and 4.55% of explained variance, respectively. The third mode, showed in Figure 7, had separated regions of the spatial patterns with negative and positive values. The negative values spread around the coast of Sumatra and Java, and along the western part of the study area. The lowest negative values were seen in Sumatra coast. On the other hand, the positive values spread in the eastern part (Figure 7a). The temporal variation from the PC showed uneven fluctuations. Furthermore, based on the PC and PSD analysis, the variation was clear in being shown as annual (356 days) and semi-annual (182 days) periodicities (Figure 7b, 7c). Subsequently, the fourth mode, presented in Figure 8, showed the spatial pattern with dominant positive values along west coast of Sumatra, spread to offshore heading to the coast of Australia. The negative values were seen in two spots, one around Cilacap coast until Nusa Tenggara, and another one along southwest part of the study area (Figure 8a). The temporal pattern from the PC and PSD showed the dominant fluctuations were annual (356 days) and inter-annual (1024 days) periodicities (Figure 8b, 8c).

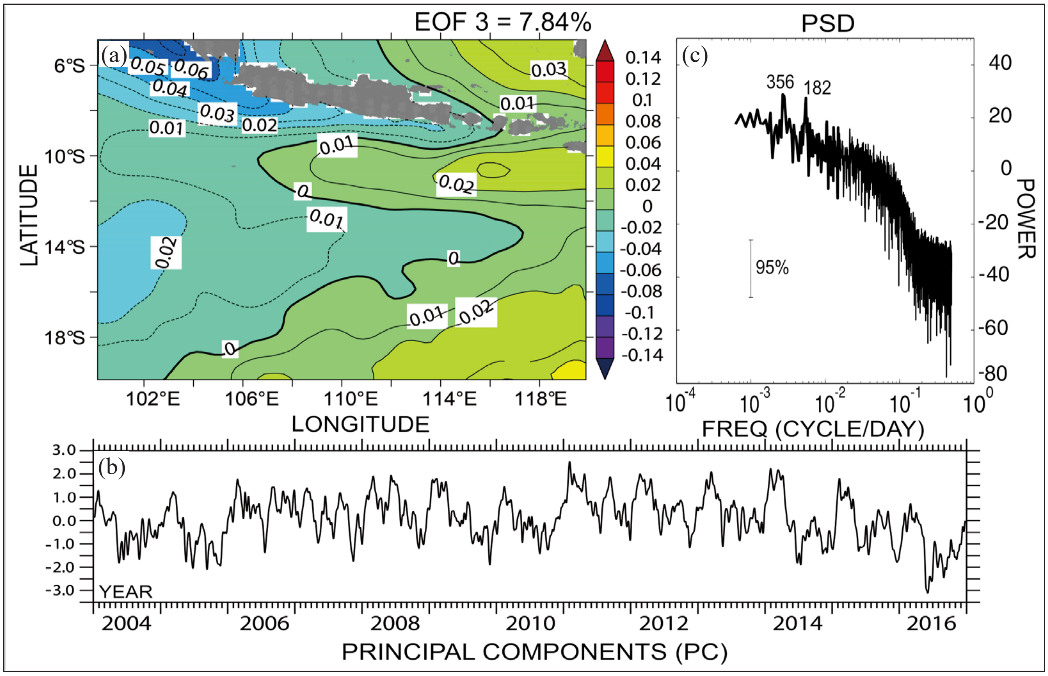


Figure 7. EOF analysis from SLA in the third mode (a) spatial pattern (b) PC graph (c) PSD analysis with 95% of confidence interval

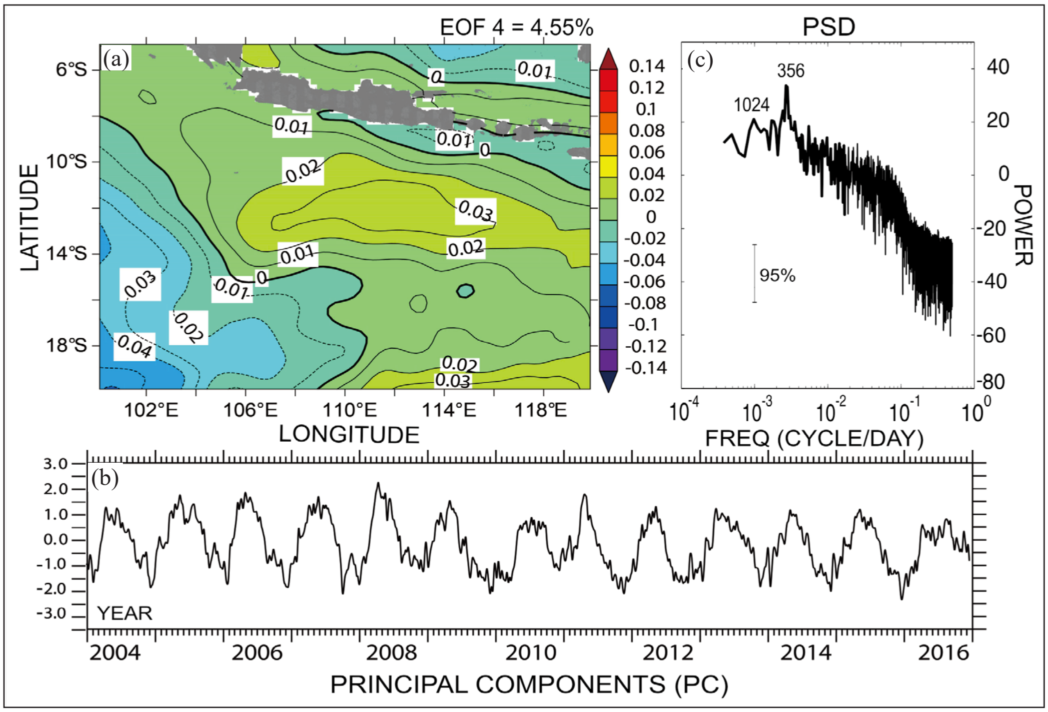


Figure 8. EOF analysis from SLA in the fourth mode (a) spatial pattern (b) PC graph (c) PSD analysis with 95% of confidence interval

The spatial patterns of geostrophic zonal component were interpreted from EOF analysis in the EIO region. Figure 9 showed EOF for geostrophic zonal component to present spatial and temporal patterns of the geostrophic zonal component in the first mode, accounting for 22.27% of the explained variance. For spatial pattern, there was a positive values around the south coast of Java until 12.5°S with the highest positive values lied along west coast of Sumatra, south coast of Java coast until Nusa Tenggara. Right below 12.5°S until the end coverage of the study area was a small variation between negative and positive values (Figure 9a). Temporal fluctuation from PC and PSD analysis in this mode showed that the annual (365 days) periodicity was dominant, followed by semi-annual (182 days) and inter-annual (1024 days) periodicity (Figure 9b, 9c).

The Second EOF mode in zonal geostrophic component was presented in Figure 10, accounting for 6.89% of the explained variance. This mode showed the positive values spreading along the southern Sumatra and Java until Nusa Tenggara coasts with the maximum centered positive values around 106°E-112°E and 8°S-8.5°S. The spatial pattern also showed negative values around 10°S-13°S and the lowest value were centered around 107°E-115°E 11°S-13°S. The contour located below 13°S did not show significant

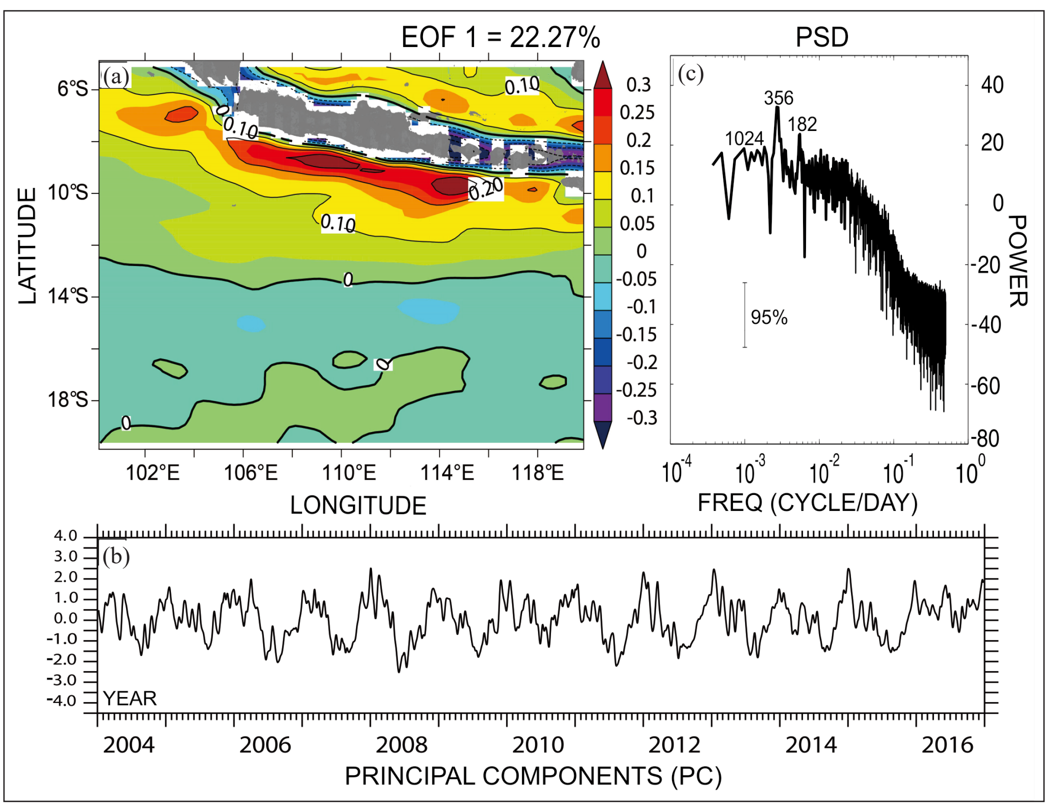


Figure 9. EOF analysis from geostrophic zonal component in the first mode (a) spatial pattern (b) PC graph (c) PSD analysis with 95% of confidence interval

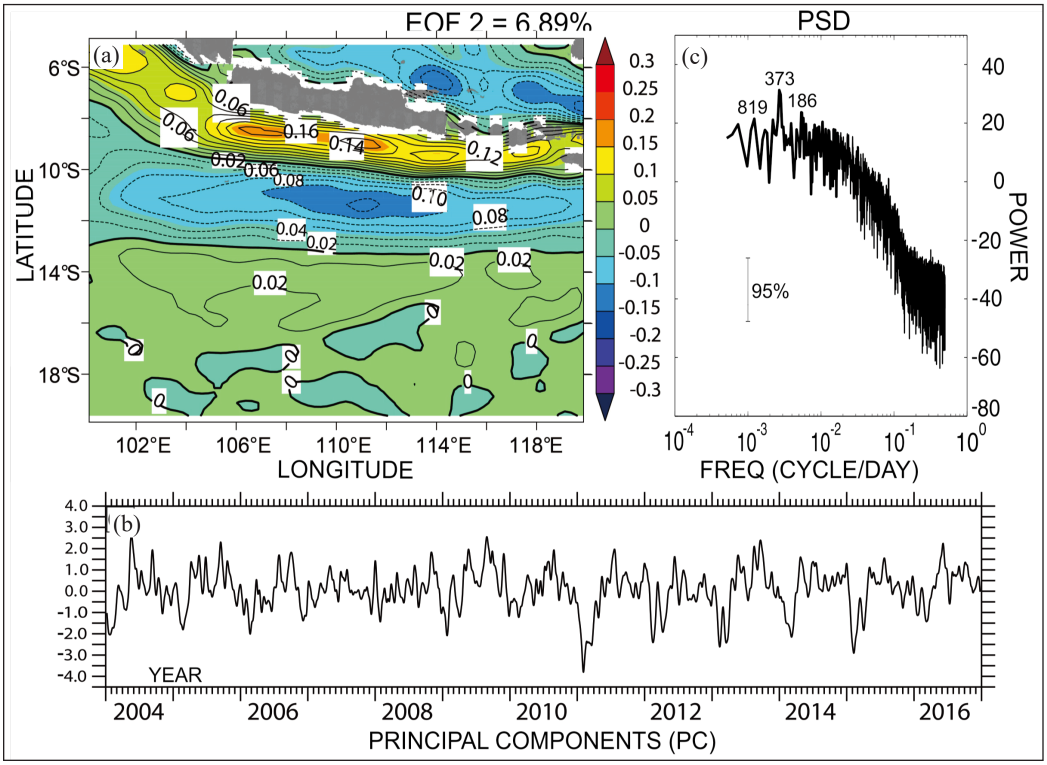


Figure 10. EOF analysis from geostrophic zonal component in the second mode (a) spatial pattern (b) PC graph (c) PSD analysis with 95% of confidence interval

fluctuations (Figure 10a). Temporal variation showed from the PC were annual (373 days), semi-annual (186 days) and inter-annual (819 days) periodicities (Figure 10b, 10c).

The third and fourth mode were minor mode regarding to small number of explained variance. The third mode accounted for 3.6% of the explained variance (Figure 11). The spatial pattern was presented in Figure 11a, showed positive values, seen around the west coast of Sumatra to the south coast of Java, which was almost similar to the second mode. Meanwhile, the negative values were seen around 6°S-10.5°S. Moreover, the rest of the eigenvector values spread along the offshore part of the study area, showing few variations indicated with the small values of positive and negative spatial patterns. The temporal variation from the PC and PSD were observed to be annual (373 days) and semi-annual (186 days) periodicity (Figure 11a, 11b). The fourth mode, accounting for 2.89% of the explained variance (Figure 12), showing the spatial pattern with negative eigenvector around the coast of Sumatra up to 110°E. The negative values were also seen around 100°E-116°E and 10°S-13°S (Figure 12a). Meanwhile, the positive eigenvector values were shown at the eastern part of south Java coast, forming centralization with positive values around 110.5°E -116°E and 9°S-11°S. The temporal variation from the PC and PSD analysis showed annual (372 days) and semi-annual (163 days) periodicities (12b, 12c).

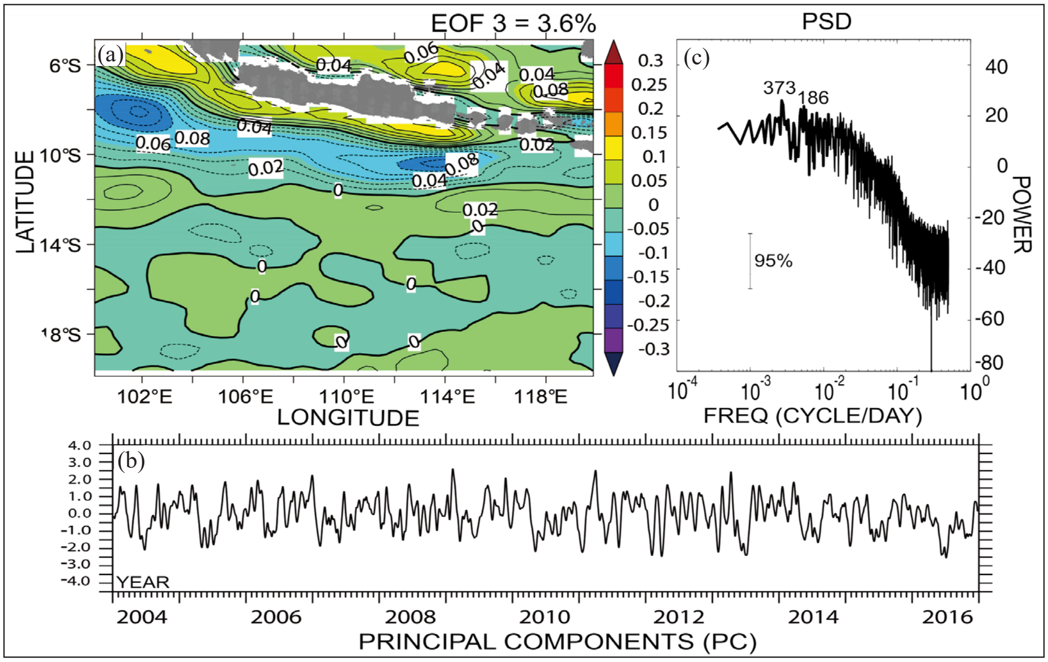


Figure 11. EOF analysis from geostrophic zonal component in the third mode (a) spatial pattern (b) PC graph (c) PSD analysis with 95% of confidence interval

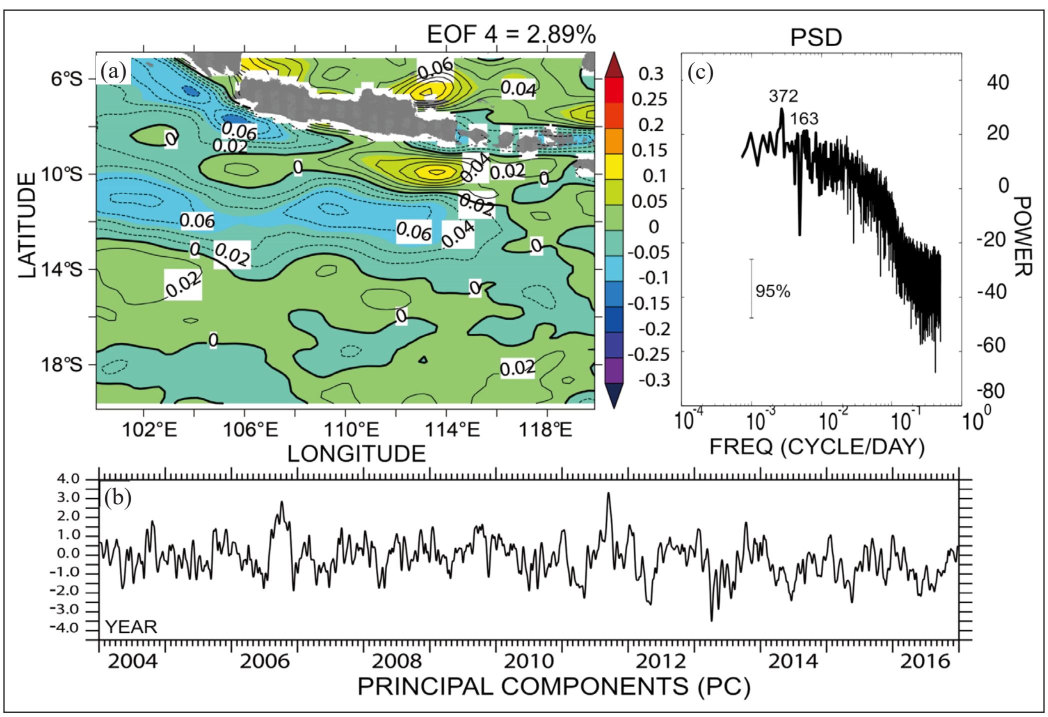


Figure 12. EOF analysis from geostrophic zonal component in the fourth mode (a) spatial pattern (b) PC graph (c) PSD analysis with 95% of confidence interval

The highest standard deviation occurred near the coast of Java both in the SLA and the zonal geostrophic component data (Figures 2b, 2c). Furthermore, there was also a high standard deviation in the area far off Java island. The higher values of standard deviation near the south coast of Java, both in SLA and geostrophic zonal component, might indicate high variability in the area. The highest contour values seen in 111°E-116°E and 8°S-9°S could be affected by Kelvin wave propagation pathway (Clarke & Liu, 1993; Sprintall et al., 2000). The Kelvin wave moves from the equator of the Indian Ocean and causes downwelling near the southern coast of Java. This area is also the pathway of Southeast and northwest monsoon so the area experienced positive-negative anomalies (Susanto et al., 2001). The contour of the minimum standard deviation was seen from the west and east. It might be associated with inter-annual IOD and ENSO phenomena. A more complete descriptions of the area were explained through the results of the EOF analysis.

The results of EOF analysis of SLA from the first to the fourth modes showed variability with semi-annual, annual and inter-annual periodicity. The sum of explained variance was 79.68%. The EOF of SLA in the first mode presented high maximum values along the west coast of Sumatra, the south coast of Java and Nusa Tenggara (Figure 5a). Furthermore, the temporal variation revealed the dominant energy, oscillating in semi-annual, annual and inter-annual periodicity. Special attention was required to analyze the spatial and temporal variations of the SLA in the first mode (Figure 5b,5c). The high positive phase along Sumatra and Java coast might be associated with CTKW that happened semiannually (Sprintall et al., 2000; Syamsudin, 2004). Agreeing with this result, Potemra and Schneider (2007) pointed out that semi-annual Kelvin wave was the large effect of westerly wind bursts during monsoon transition in the Indian Ocean. Sprintall et al. (2000) mentioned that during CTKW, SLA rose along the coast of Sumatra and Java.

Moreover, the temporal variability from PC showed some peaks in April and May, which matched with time series of SLA analysis carried out by Sprintall et al. (2000). The SLA time-series analysis showed that during strong Kelvin wave, the propagation in the monsoon transition period (April-May) was related to the westerly wind bursts. The high SLA signal was associated with equatorial downwelling Kelvin wave propagating eastward, forced by semi-annual westerly wind bursts. Although Kelvin wave propagates semiannually, Wyrtki (1973) said that it would be weaker in the monsoon transition in the October-November period. Syamsudin (2004) pointed out that during 1993-2001, the occurences of Kelvin wave happened either in April-May or November-December. In our results, the occurences of Kelvin wave mostly happened in April-May and November-December, and the weaker November-December Kelvin wave occurred in 2004, 2006, 2009, 2012, 2013 and 2015. Other signals revealed in the first mode's PSD were annual and inter-annual. Although the annual periodicity had been detected in the first mode, based on spatial and temporal variation, it was more significant in the second mode. So, the annual variability of SLA was presented in the second mode.

The SLA variability from second mode was presented in Figure 6. The spatial pattern showed negative values along west coast of Sumatra and south coast of Java while positive values spread in the offshore of southern Java Sea (Figure 6a). The temporal variation also showed the dominant fluctuation's peaks amplitude occurring in July-August-September in the spanning of 2004-2016, and the strongest density energy signal was observed to have annual variation (Figure 6b,6c). Based on that evidence, it could be predicted that the oscillations happening in this mode was influenced by annual monsoon generating upwelling signal. Using the standard deviation of the monthly sea surface temperature, Susanto et al. (2001) revealed annual upwelling along the south coast of Java and west coast of Sumatra. Annual upwelling generated by southeast monsoon happened from June to October. This phenomenon affects colder SST and lower sea level. Another phenomenon that could be seen from this mode was downwelling, happened during northwest monsoon in December-Januari-February, as appeared in the PC's negative peaks.

Inter-annual variability of SLA was seen in the first mode based on temporal variation from PC and correlation coefficient results between the temporal variation of the first mode and IOD index. The correlation coefficient between the IOD index and the first mode of PC from SLA data was -0.41, meaning that there was an inverse correlation between those two variables. During 2004-2016, there were positive (2006, 2012 and 2015) and negative (2010, 2014 and 2016) IOD indexes. In our data, we found the lowest negative anomaly of SLA occurred in 2006. It linked to 2006 positive IOD that consistent with previous study (Horii et al., 2008; Lumban-Gaol et al., 2015). The negative anomaly of SLA in the EIO is occurred when the easterly wind formed in the equator of Indian Ocean, allowing warm pool to move to Africa. Otherwise, we also found the positive anomaly of SLA during negative IOD in 2016. In accordance with Lim & Hendon (2017), the SLA was higher during strong negative IOD in 2016 because the westerly wind intensify, allowing warm pool to concentrate near Indonesia and Australia. Saji et al. (1999) stated that the IOD phenomenon had a powerful influence on the Indian Ocean. Furhermore, Sprintall et al. (2009) found that in 2006, a positive IOD phase coincided by a strong El Niño. In addition, Lim & Hendon (2017) stated, in 2016, there was a strong negative IOD coincided by weak La Niña.

Inter-annual variability in SLA also occurred in ENSO that happened in Pacific Ocean. We found that the variability of SLA experienced a dramatic change when entering the ENSO period. For instance, La-Niña period in 2011, it was seen that the value of PC from SLA data in the first mode showed the highest positive anomaly. The positive anomalies were related to warm pool which was intensively pushed into Indonesia area due to trade wind in equator Pacific Ocean that blew westward during the La-Niña period (Lim & Hendon, 2017). On the other hand, the SLA experienced negative anomaly in 2015, coinciding with El-Niño period. In El-Niño period, the weakening of the trade wind caused the warm pool to move back into South America, causing SLA around Indonesian seas to be lower. It was consistent with previous studies (Cao et al., 2018; Syamsuddin et al., 2013).

Another variability with lesser contribution showed in the third mode. The spatial pattern in the third mode, accounting for 7.84% of explained variance, got influence from the west area, which might be influenced by the Indian Ocean with negative anomalies. Furthermore, the pattern at the east area might be influenced by the Pacific Ocean with positive anomalies. The temporal variation from the PC graph showed the majority of positive peaks happened in February, March, June, and September. While the negative peaks happened in June, July, and October, November, and December indicated by annual and semi-annual periodicity.

The fourth mode, accounting for 4.55% of explained variance, indicated a different phase of spatial pattern, but not significant. Almost all areas were influenced by positive anomalies. Only two spots were indicated by negative anomalies, around Cilacap and in the southwest of the study area, off the coast. The positive peaks seen from PCs were in April and May, except 2007 that occurred in June, 2010 that occurred in July, and 2016 that occurred in September. While the negative peaks mostly happened in October, November, and December. Except in 2006 and 2016 which was in January. The dominant periodicity that occurred in this area was annual variability.

The spatio-temporal variability of geostrophic zonal component was described in the highest four-modes, accounted for 36.65% of cummulative explained variance. It was far from the threshold of determining how many modes would be used provided by Jolliffe (2002), but it was not possible to use more modes, because the explained variance percentage was getting smaller for the fifth mode and forward. This means that it was not possible to explain spatial and temporal variability from a small percentage of explained variance (<2.89%). So only the highest four-modes were included in this study to make it consistent with SLA modes.

The spatial pattern of the geostrophic zonal component in the three leading EOF showed positive values around Sumatra and Java coast as showed in Figure 9a, 10a, and 11a, respectively. In the spatial pattern of the first mode, there was positive values along the west coast of Sumatra-the south coast of Java, centralized around the south coast of Java. Referring to the annual period of the variability, it was suspected to be annual upwelling. It occurred during the southeast monsoon that flowed westward, removing light surface water and generating upwelling along Java coast (Fieux et al., 1996). Hence, water circulation flows westward. During monsoon transition, westerly wind forces from the equatorial Indian Ocean, generating Kelvin wave with the semi-annual period (Sprintall et al., 2000). The semi-annual signal was shown in several years in the PC, for example, in April 2004 and May 2005. Besides that, the inter-annual variability in this mode leads to some arguments. Chen, et al. (2016) argued that inter-annual variability in the equatorial EIO is upwelling and closely associated with IOD, while a few were associated with the ENSO event. Regarding the ENSO, it was reported that the ITF was higher during La Niña than El-Niño (Meyers, 1996). On the other hand, Sprintall et al., (2009) found that the surface ITF anomaly was strong during positive IOD coinciding with strong El-Niño in 2006.

The second mode showed positive values along Sumatra and Java coast, and the typical pattern was considered as the SJC pathway (Figure 10a). The SJC is generated by local wind stress along the coast of Sumatra and Java and part of the easterly semi-annual equatorial jet reflection generated in central Indian Ocean during monsoon transition (Clarke & Liu, 1993; Meyers et al., 1995). The maximum eastward SJC flows in May and November (Bray et al., 1997). The negative values, seen around 10°S-13°S, was the same location with the SEC region, that was consistent with previous study (Feng & Wijffels, 2002). The SEC annually flows from ITF, where it is transported to the EIO that flows westward, carrying a mass of water with low salinity. The highest intensity of the geostrophic component in the SEC was in July-November, and the weak energy transport happened when the eastward SJC strengthened (Bray et al., 1997; Feng & Wijffels, 2002; Meyers et al., 1995).

The third and fourth modes contributed very little variability, with 3.6% and 2.89% of explained variance. In the third mode, there was a spatial pattern with positive anomalies near the west coast of Sumatra and South of Java, while the negative anomalies were around 6°S-10.5°S. The spatial pattern was similar to the second mode, it might be associated with SJC and SEC propagation path. The temporal variation of the PC showed positive peaks in February, March, September, November, and December with annual, semi-annual and seasonal periodicity. Whereas in the fourth mode, although it had positive and negative anomalies with a number of centralization, the variation was not significant. The temporal period oscillated in January, February, September, and October, and negative peaks were in May, April, and July. The variability oscillated in annual, inter-annual and semi-annual periodicity.

CONCLUSION

The investigation result of SLA and geostrophic zonal component showed great variations near the south coast of Java. The spatial pattern of SLA in the first mode was dominated by positive eigenvector with the high value around Sumatra and Java coast, while the second mode showed negative values around Sumatra and Java coast, and positive values along the offshore region, while the third and the fourth mode did not show a substantial spatial variability. The temporal variability results showed semi-annual, annual and inter-annual periodicity. Based on the spatial and temporal patterns of SLA data, it was suspected that SLA variability that occurred in the EIO region might be associated with semi-annual Kelvin wave, annual upwelling, and inter-annual IOD and ENSO phenomena.

Furthermore, the EOF analysis for geostrophic zonal component in the first mode showed that there were positive values around Sumatra and Java coast, while negative value spread along the offshore region. In the second mode, the spatial pattern showed positive values near the coastal area, and negative values around 10°S-13°S, with annual, semi-annual and inter-annual periodicity. In the third and the fourth mode, the spatial pattern showed low variability. The dominant temporal variation revealed annual, semi-annual and inter-annual periodicity. It was also suspected that spatial and temporal variability in the geostrophic zonal component was supposed to be upwelling, Kelvin wave, and the ITF during IOD and ENSO. It was noticed that some patterns could be associated with the SJC and SEC signals.

ACKNOWLEDGMENT

This study has been conducted using E.U Copernicus Marine Service Information. The authors would also like to express their special gratitude to the colleagues at Marine Technology major, the Oceanographic Data Processing Laboratory and colleagues at the Oceanographic Physics Laboratory of IPB University for their suggestions and input for the sake of accomplishing this research.

REFERENCES

Atmadipoera, A. S., Horhoruw, S. M., Purba, M., & Nugroho, D. Y. (2016). Spatial and temporal variation of Indonesian Throughflow in the Makassar Strait. *Ilmu Dan Teknologi Kelautan Tropis*, *8*(1), 299-320.

AVISO. (2016). *SSALTO/DUACS User Handbook : MSLA and (M)ADT Near-Real Time and Delayed Time Products*. France (FR): CNES.

Björnsson, H., & Venegas, S. (1999). *A Manual for EOF and SVD analyses of climatic data*. Montreal, Canada: McGill University.

Bray, N. A., Wijffels, S. E., Chong, J. C., Fieux, M., Hautala, S., Meyers, G., & Morawitz, M. L. (1997). Characteristics of the Indo-Pacific throughflow in the eastern Indian Ocean. *Geophysical Research Letters*, *24*(21), 2569-2572.

Cao, G. J., He, Y. J., Wei, Z. X., & Xu, T. F. (2018). Interannual modulation of intraseasonal sea level variability along the southern coast of Java. *Science China Earth Sciences*, *61*(1), 1-12.

Chen, G., Han, W., Li, Y., & Wang, D. (2016). Interannual variability of Equatorial Eastern Indian Ocean upwelling : Local versus remote forcing. *Journal of Phisical Oceanography*, *46*, 789-807.

Clarke, A. J., & Liu, X. (1993). Observations and dynamical of semiannual and annual sea levels near the eastern equatorial indian ocean boundary. *Journal of Physical Oceanography*, *23*, 386-399.

Escudier, P., Couhert, A., Mercier, F., Mallet, A., Thibaut, P., Tran, N., ... & Ablain, M. (2017). *S*atellite altimeter over ocean and land surfaces. In D. Stammer & A. Cazenave (Eds.), *Satellite altimetry over oceans and land surfaces* (pp. 1-70). Boca Raton, Florida: CRC Press.

Fang, W., Guo, J., Shi, P., & Mao, Q. (2006). Low frequency variability of South China Sea surface circulation from 11 years of satellite altimeter data. *Geophysical Research Letters*, *33*(22), 1-5.

Feng, M., & Wijffels, S. (2002). Intraseasonal variability in the SEC of the East Indian ocean. *Journal of Geophysical Research*, *32*, 265-277.

Fieux, M., Molcard, R., & Ilahude, A. G. (1996). Geostrophic transport of the Pacific-Indian Oceans throughflow. *Journal of Geophysical Research*, *101*(95), 12421-12432.

Fitrianah, D., Hidayanto, A. N., Lumban-Gaol, J., Fahmi, H., & Arymurthy, A. M. (2016). A spatio-temporal data-mining approach for identification of potential fishing zones based on oceanographic characteristics in the Eastern Indian Ocean. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *9*(8), 3720-3728.

Fu, L. L., & Cazenave, A. (2001). *Satellite altimetry and earth sciences*. San Diego, California: Academic Press.

Fu, L. L., & Chelton, D. B. (1985). Observing large-scale temporal variability of ocean currents by satellite altimetry: With application to the Antarctic circumpolar currents. *Journal of Geophysical Research*, *90*(C3), 4721-4739.

Gill, A. E. (1982). *Atmosphere-ocean dynamics*. New York, NY: Academic Press.

Gordon, A., & Kamenkovich, V. (2010). “Modeling and observing the Indonesian Throughflow ” a special issue of dynamics of atmosphere and ocean. *Dynamic of Atmospheres and Ocean*, *50*, 113-114.

Hannachi, A. (2004). *A Primer for EOF analysis of climate data* (pp. 1-33). England, UK: University of Reading.

Hannachi, A., Jolliffe, I. T., & Stephenson, D. B. (2007). Empirical orthogonal functions and related techniques in atmospheric science: A review. *International Journal of Climatology*, *27*(December 2007), 1549-1555.

Horii, T., Hase, H., Ueki, I., & Masumoto, Y. (2008). Oceanic precondition and evolution of the 2006 Indian Ocean dipole. *Geophysical Research Letters*, *35*(3), 1-6.

Izumo, T., Vialard, J., Lengaigne, M., de Boyer Montegut, C., Behera, S. K., Luo, J. J., … & Yamagata, T. (2010). Influence of the state of the Indian Ocean Dipole on the following years El Nĩno. *Nature Geoscience*, *3*(3), 168-172.

Jolliffe, I. T. (2002). *Principal component analysis, Second Edition*. New York, NY: Springer-Verlag.

Joseph, A. (2014). *Measuring ocean currents: Tools, technologies, and data* (1st Ed.). Waltham, MA: Elsevier.

Le Traon, P. Y. (2013). From satellite altimetry to Argo and operational oceanography : three revolutions in oceanography. *Ocean Science*, *9*(5), 901-915.

Lim, E., & Hendon, H. H. (2017). Causes and predictability of the negative Indian Ocean Dipole and its impact on La Niña During 2016. *Nature*, (June), 1-11.

Liu, Y., & Weisberg, R. H. (2007). Ocean currents and sea surface heights estimated across the West Florida Shelf. *Journal of Physical Oceanography*, *37*(6), 1697-1713.

Lumban-Gaol, J., Leben, R. R., Vignudelli, S., Mahapatra, K., Okada, Y., Nababan, B., … & Syahdan, M. (2015). Variability of satellite-derived sea surface height anomaly, and its relationship with Bigeye tuna (Thunnus obesus) catch in the Eastern Indian Ocean. *European Journal of Remote Sensing*, *48*, 465-477.

Mansawan, A. A., Lumban-Gaol, J., & Panjaitan, J. P. (2016). Variation and trend of sea level derived from altimetry satellite and tide gauge in cilacap and benoa coastal areas. *International Journal of Remote Sensing and Earth Sciences*, *13*(1), 59-66.

Meyers, G. (1996). Variation of Indonesian throughflow and the El Nino - Southern Oscillation. *Science China Earth Sciences*, *101*(95), 12255-12263.

Meyers, G., Bailey, R. J., & Worby, A. P. (1995). Geostrophic transport of Indonesian throughflow. *Deep-Sea Research Part I*, *42*(7), 1163-1174.

Pond, S., & Pickard, G. L. (1983). *Introductory dynamical oceanography* (2nd Ed.). Oxford, England: Butterworth-Heineann Ltd.

Potemra, J. T., & Lukas, R. (1999). Seasonal to interannual modes of sea level variability in the western Pacific and eastern Indian oceans. *Geophysical Research Letters*, *26*(3), 365-368.

Potemra, J. T., & Schneider, N. (2007). Interannual variations of the Indonesian throughflow. *Journal of Geophysical Research: Oceans, 112*(C5), 1-13.

Rahmawitri, H., Atmadipoera, A. S., & Sukoraharjo, S. S. (2016). Circulation and current variability in the Sunda Strait waters. *Jurnal Kelautan Nasional*, *11*(3), 141-157.

Rosmorduc, V., Benveniste, J., Bronner, E., Dinardo, S., Lauret, O., Maheu, C., … & Picot, N. (2011). *Radar Altimetry Tutorial*. Howard Place, South Africa: CSL Publisher.

Saji, N. H., Goswami, B. N., Vinayachandran, P. N., & Yamagata, T. (1999). A dipole mode in the tropical Indian Ocean. *Nature*, *401*, 360-363.

Schiller, A., Wijffles, S., Sprintall, J., Molcard, R., & Oke, P. (2009). Pathways of intraseasonal variability in the Indonesian throughflow region. *Dynamics of Atmospheres and Oceans, 50*(2), 174-200.

Schott, F. A., Xie, S. P., & McCreary-Jr, J. P. (2009). Indian Ocean circulation and climate variability. *Reviews of Geosphisics*, *47*(RG1002), 1-46.

Song, Q., Gordon, A., & Visbeck, M. (2004). Spreading of the Indonesian throughflow in the Indian Ocean. *Journal of Physical Oceanography*, *34*, 772-792.

Sprintall, J., Gordon, A., Murtugudde, R., & Susanto, R. D. (2000). A semiannual Indian Ocean forced Kelvin wave observed in the Indonesian seas in May 1997. *Journal of Geophysical Research*, *105*, 17217-17230.

Sprintall, J., Wijffels, S. E., Molcard, R., & Jaya, I. (2009). Direct estimates of the Indonesian Throughflow entering the Indian Ocean : 2004 – 2006. *Journal of Geophysical Research*, *114*(C07001), 1-19.

Sprintall, J., Wijffels, S., Molcard, R., & Jaya, I. (2010). Direct evidence of the South Java Current system in Ombai Strait. *Dynamics of Atmospheres and Oceans*, *50*(2), 140-156.

Stewart, R. H. (2008). *Introduction to Physical Oceanography*. Texas, USA: Texas A & M University.

Susanto, R. D., Ffield, A., Gordon, A., & Adi, T. R. (2012). Variability of Indonesian throughflow within Makassar Strait , 2004-2009. *Journal of Geophysical Research*, *117*(C09013), 2004-2009.

Susanto, R. D., Gordon, A., & Zheng, Q. (2001). Upwelling along the coasts of Java and Sumatra and its relation to ENSO. *Geophysical Research Letters*, *28*(8), 1599-1602.

Susanto, R. D., & Song, Y. (2015). Indonesian throughflow proxy from satellite altimeters and gravimeters. *Journal of Geophysical Research : Oceans*, *120*, 2844-2855.

Syah, A. F., Saitoh, S. I., Alabia, I. D., & Hirawake, T. (2016). Predicting potential fishing zones for pacific saury (Cololabis saira) with maximum entropy models and remotely sensed data. *Fishery Bulletin*, *114*(3), 330-342.

Syamsuddin, M. L., Saitoh, S. I., Hirawake, T., Samsul, B., & Harto, A. B. (2013). Effects of El Niño-Southern Oscillation events on catches of Bigeye Tuna (Thunnus obesus) in the eastern Indian Ocean off Java. *Fishery Bulletin*, *111*(2), 175-188.

Syamsudin, F. (2004). Semiannual Kelvin Waves Propagation along the South Coast of Sumatra-Jawa-Bali and the Lesser Sunda Islands observed by TOPEX / POSEIDON and ERS-1 / 2 Satellites. *Journal of Engineering and Technological Sciences*, *36*(2), 133-139.

Syamsudin, F., & Kaneko, A. (2013). Ocean variability along the southern coast of Java and Lesser Sunda Islands. *Journal of Oceanography, 69*(5), 557-570.

Thomson, R. E., & Emery, W. J. (2014). *Data analysis method in physical oceanography* (3rd Ed.). Amsterdam, The Netherlands: Elsevier.

Tomczak, M., & Godfrey, J. S. (2001). *Regional oceanography: An introduction* (1.0). London, UK: Elsevier.

Trauth, M. H. (2014). *MATLAB ® Recipes for earth sciences* (4th Ed.). Berlin, Germany: Springer.

Wyrtki, K. (1961). *NAGA REPORT, Scientific result of marine investigaitons of the South China Sea and the Gulf of Thailand 1959-1961* (Vol. 2). University of California, Scripps Institution of Oceanography, California.

Wyrtki, K. (1973). Physical oceanography of the Indian Ocean. In B. Zeitzschel & S. A. Gerlach (Eds.), *The biology of the Indian Ocean* (Vol. 3, pp. 18-23). Berlin, Germany: Springer.

Pertanika J. Sci. & Technol. 27 (4): 2305 - 2315 (2019)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Strength Assessment of Mining Soil Treated with Steel Slag as Liner Material for Retention Pond

Noratiqah Masri[1], Faradiella Mohd Kusin[1,2]*, Nik Norsyahariati Nik Daud[3] and Hasfalina Che Man[4]

[1]Department of Environmental Sciences, Faculty of Environmental Studies, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia
[2]Environmental Forensics Research Unit (ENFORCE), Faculty of Environmental Studies, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia
[3]Department of Civil Engineering, Faculty of Engineering, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia
[4]Department of Biological and Agricultural Engineering, Faculty of Engineering, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia

ABSTRACT

Clay soil has always been associated with low shear strength and high compression behavior due to the high content of organic matter. The limited amounts of clay available onsite and acid mine drainage (AMD) problems have necessitated the continuous search for the treatment technology potentials. Mining soils, obtained from Selinsing Gold Mine in Raub, Pahang were evaluated to determine their suitability for use as mining soil and steel slag mixtures as compacted retention pond liners for AMD treatment. The studied samples were subjected to classification, compaction, permeability and strength tests. The results indicated that the index properties of the samples met the minimum requirements for use as liners. The compaction test showed that the maximum dry density (MDD) and optimum moisture content (OMC) decreased and increased, respectively, for all studied samples. At OMC, hydraulic conductivities of the compacted soil-steel slag were in the order of $\leq 10^{-9}$ m/s. The results from unconfined compression strength (UCS) tests gave values of 204 kN/m^2 and 61° for soil cohesion and soil internal friction angle, respectively. Furthermore, the influence of steel slag treatment on strength properties has generally shown an improvement of up to 15% steel slag which gives the acceptable results of stress-strain in respect of its usability as liner material.

Keywords: Hydraulic conductivity, liner, mining soil, steel slag, strength parameters

ARTICLE INFO

Article history:
Received: 03 December 2018
Accepted: 30 January 2019
Published: 21 October 2019

E-mail addresses:
atiqahmasri93@gmail.com (Noratiqah Masri)
faradiella@upm.edu.my (Faradiella Mohd Kusin)
niknor@upm.edu.my (Nik Norsyahariati Nik Daud)
hasfalina@upm.edu.my (Hasfalina Che Man)
* Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

INTRODUCTION

Mining industry continues to play an important role in providing basic resource to the construction and industrial sectors, contributing to the economic development of many countries around the world including Malaysia. However, certain types of mine tailings will oxidize with oxygen and water, which results in the formation of acid mine drainage (AMD) (Fashola et al., 2016; Kusin et al., 2018). If not properly managed, the disposal of acid generating mining wastes and dissolution of metals will cause pollution that contaminates surface and groundwater (Kusin et al., 2017; Masri et al., 2017). Therefore, the mining industry relies on diverse types of liners to minimize the effluent migration from tailings impoundments to surface and/or groundwater.

There have been numerous research and procedures that focused on the mitigation measurements of AMD. Previous studies have shown examples of passive treatment in the remediation of acidic water and have been practically applied in the mining site as the systems are environmentally friendly and cost-effective (Fashola et al., 2016; Miguel et al., 2015; Muhamad et al., 2017; Molahid et al., 2018; Zahar et al., 2015;). The use of retention pond as a treatment system represents an alternative to chemical treatment techniques. The systems are constructed directly on mine tailings impoundments which have been discussed by Lagos and Geo (2011). In these systems, liner materials have become critical components particularly in the design and performance of mining treatment systems.

Commonly, the liner materials used are, for example, clay (Rowe et al., 2004; Wagner & Schnatmeyer, 2002) or geosynthetic clay (Wagner & Schnatmeyer, 2002). The characteristics of used materials can strongly influence the performance of liner. In order to reduce the cost of construction due to lack of availability of the clay onsite, utilization of possible alternative materials has been proposed, such as soil-POFA mixtures (Nik Daud& Mohammed, 2014), shale-clay mixtures (Li et al., 2017), soil from ore treatment (Miguel et al., 2015), and steel slag (Herrmann et al., 2010). Instead of using clay soil as a liner, mining soil that possesses suitable geotechnical characteristics can be considered as an alternative material.

A recent study by Rowe et al. (2004) outlined that liners could either be artificially synthesized seals or compacted clay liners and should have hydraulic conductivity of less than or equal 10^{-9} m/s. The specified maximum hydraulic conductivity also conforms to the United States Environmental Protection Agency (USEPA, 1989). According to Benson et al. (1994), Rowe et al. (2004) and Nik Daud and Mohammed (2014), soil samples that have the characteristic as presented in Table 1 mostly meet the requirements for liner materials. The selection and performance standards for assessing the suitability of the materials intended for hydraulic barriers in retention pond are low hydraulic conductivity and adequate shear strength.

Table 1
Properties requirements for liner materials

Properties	Limiting value		
	Benson et al. (1994)	**Rowe et al. (2004)**	**Nik Daud et al. (2014)**
Gravels (%)	-	< 50	-
Clay (%)	> 15	-	-
Fines (%)	> 30	> 15 – 20	≥ 20 – 30
Plasticity index (%)	> 7	> 7	≥ 7
Liquid limit (%)	> 20	-	> 20
Hydraulic conductivity (m/s)	$\leq 1 \times 10^{-9}$	$\leq 1 \times 10^{-9}$	$\leq 1 \times 10^{-9}$

Thus, this paper provides information on the properties of mining soil and steel slag mixture as an improved retention pond liner material. This study aims to assess the compressive strength, obtained by the mining soil-steel slag mixtures.

MATERIALS AND METHODS

Materials

Native mining soil samples, exploited in this study were obtained from Selinsing Gold Mine area in Raub, Pahang, Malaysia (GPS coordinate: $4^0$15’42.86” N $101^0$46’52.75” E) (Figure 1). Three samples of soils from the same location were collected using a hammer, digger and shovels about 20-50 cm from the surface soils. Disturbed soil samples were collected and sealed in a polyethylene bag to retain the soil moisture in accordance with BS 1377-1 (Road Engineering Standards Policy Committee, 1990).

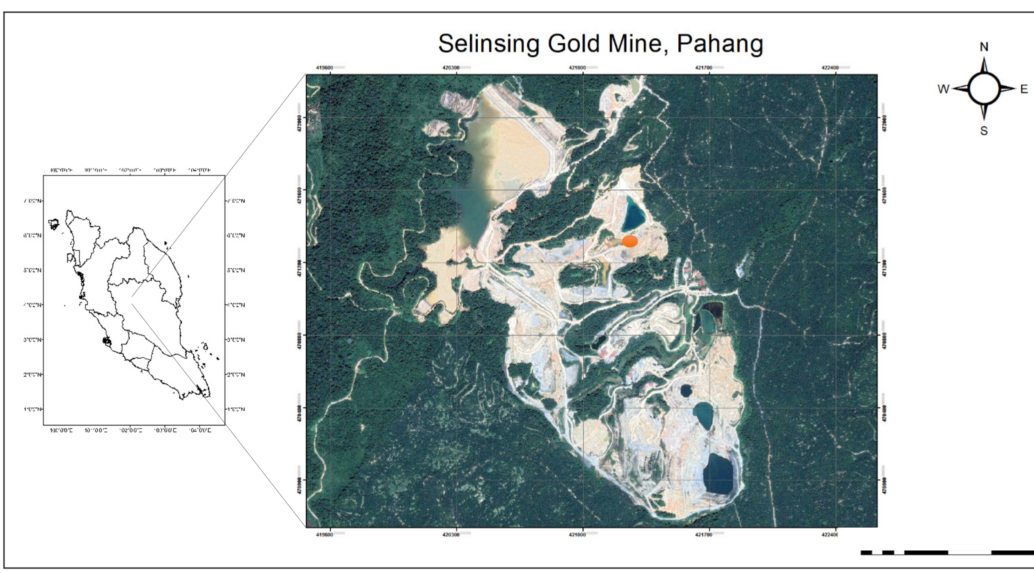


Figure 1. Location map of mining soil study area

The soil was characterized according to its typical physical properties from a laboratory test (Table 2). Soil specimens were prepared based on their average moisture content level before they were dry-mixed with steel slag at designated optimum proportion.

Table 2
Physical properties of untreated mining soil (Masri et al., 2017)

Properties	**Mine soil**
pH	4.4
Mechanical analysis	
Sand (%)	1.19
Silt (%)	91.05
Clay (%)	7.81
Moisture content (%)	25.39
Specific gravity of soil (g cm^{-3})	2.17
Specific Surface Area (m^2/gm)	50.00
Atterberg limits	
Liquid limit (LL)	47
Plastic limit (PL)	26.45
Plasticity index (PI)	20.55
Mineralogical composition	Kaolinite, Illite
Chemical composition	
SiO_2 (%)	75.26
Al_2O_3 (%)	19.25
K_2O (%)	4.11
Fe_2O_3 (%)	-

Meanwhile, steel slag samples were obtainable as by-products in metallurgical industry in the form of fine aggregates. The steel slag examined in this study (Table 3) was provided by a steel production company in Selangor.

Table 3
Properties of steel slag (Zahar et al., 2015)

Parameter	**Value**
BET surface area, m^2/g	30.268
Pore volume, cc/g	0.028
Pore radius, Å	15.364

The collected soil samples were air-dried, crushed and pulverized into fine grained particles in order to minimize the effects of particles size on the hydraulic conductivity of the compacted soil-steel slag. As indicated by Daniel and Benson (1990), the hydraulic conductivity of the liners increased when large particles sizes presented in the compacted soil liners.

Methods

The sample classification tests (pH, grain size distribution, specific gravity and Atterberg limits), compaction tests, permeability tests and strength tests were carried out on the samples in accordance with British Standard (BS) methods (1990). The collected soil samples were categorized parallel with requirements by the Unified Soil Classification System (USCS) using results from sieve analysis, liquid limits and plastic limits test (ASTM D2487-11).

All assessments were conducted on the mining soil treated with 0, 5, 10, and 15% steel slags by mass of the soil, respectively. Moisture content was determined immediately after sampling to obtain the initial moisture content of the material in order to avoid variations in results. Afterwards, the soil samples were air-dried at room temperature because high temperature possibly will cause some alterations in composition. The larger materials which included stones were removed and particle size < 2 mm was used for the analysis in accordance with standard methods for materials (Road Engineering Standards Policy Committee, 1990).

Compaction was performed following BS 1377-4 for Standard Proctor. In order to come up with the optimum moisture content (OMC), initially, several specimens of samples were prepared and tested at different moulding water contents. Thus, the indication of maximum dry density (MDD) from the samples was set as a reference for all tests with modified soil, prepared with steel slag.

The permeability tests were done under falling head conditions using a rigid-wall compaction mold permeameter in accordance with the procedures, described in BS 1377-6 (Road Engineering Standards Policy Committee, 1990). The compacted samples were retained inside the compaction mold permeameter and initially, the water was allowed to seep through the compacted samples for at least 72 h. The tests continued until the hydraulic conductivity was constant in which at least four hydraulic conductivity values were acquired over the period of time.

Unconfined compression strength (UCS) test was performed according to BS 1377-7. The studied samples were prepared following exact water content and unit weight of Standard Proctor compaction. The UCS test is capable to yield more reliable measure of strength. This is particularly true for fissured, compacted soils, in which the confining pressure retains the specimen intact under load (Rauch et al., 2002). In this test, treated samples were initially trimmed into test specimens measuring 100 mm high by 50 mm in diameter. Incremental strain rate of 1.5 mm per minute was applied during shearing procedure (BS 1377-7).

RESULTS AND DISCUSSIONS

According to USCS (Unified Soil Classification System), the examined soil can be classified as inorganic clay with low plasticity, CL (Masri et al., 2017). The basic characteristics of the pulverized samples were presented in Table 4. The pH of the mining soil samples was low indicating strongly acidic soil condition. The mixture of mining soil-steel slag, however, produced differences in the pH of the samples ranging between 6.7 and 10.5 as steel slag was mildly alkaline. There was a decrease in values across each parameter with the increase in steel slag content.

Table 4
Geotechnical properties of the studied samples

Properties	Mining soil	5% Steel Slag	10% Steel Slag	15% Steel Slag
pH	4.4	6.7	8.8	10.5
Moisture content (%)	25.39	-	-	-
Specific gravity of soil (g cm^{-3})	2.17	2.05	1.93	1.87
Atterberg limits				
Liquid limit (LL)	47	45	41	38
Plastic limit (PL)	26.45	25.63	24.09	23.71
Plasticity index (PI)	20.55	19.37	16.91	14.29

According to Qian et al. (2002), the liquid limit and plasticity index of a soil liner should be at least 20% and ≥ 7%, respectively, because low hydraulic conductivity is attributed to higher liquid limit and plasticity indices. As a result, the soil and soil-steel slag mixture have generally met these criteria.

Compaction test were conducted to simulate the right mixture of moisture and load on a soil that would raise the density of such soil, thus, reducing soil settlement when subjected to dynamic load. The optimum moisture content of the mining soil-steel slag mixtures, obtained from the compaction curve was shown in Figure 2.

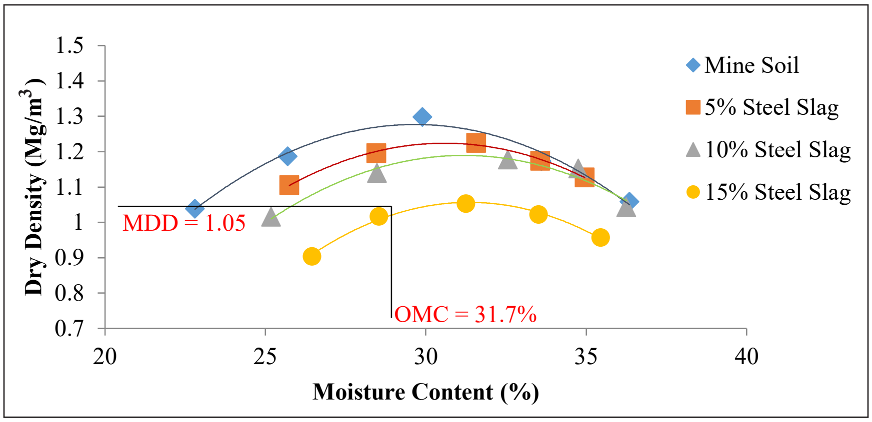


Figure 2. Compaction curves for mining soil and mining soil-steel slag mixture

Table 5 summarizes the relationship between MDD and OMC with steel slag content, clearly shows that as the steel slag content increases, OMC increases significantly, but MDD decreases. Increment in MDD with steel slag content was partly associated with steel slag's higher specific gravity compared it with that of the mining soil.

Table 5
MDD and OMC of studied samples from compaction tests

Parameter	**Mining soil**	**5% Steel Slag**	**10% Steel Slag**	**15% Steel Slag**
MDD, Mg/m^3	1.27	1.23	1.19	1.05
OMC, %	29.0	30.5	31.3	31.7

The variation findings in the MDD and OMC of the studied samples possibly associated to the differences in the plasticity. The plasticity index of 15% steel slag mixture was lower than that of the mining soil (Table 4). It was therefore expected that mining soil, treated with 15% of steel slag possess a greater affinity for water. This greater affinity for water appeared in the higher OMC and lower MDD, compared with other samples. For instance, Nik Daud and Mohammed (2014) indicated that as the plasticity index decreased, OMC increased, and the MDD decreased.

Table 6 shows that the hydraulic conductivity values generally decrease with higher steel slag content and they can be categorized as very low (Nik Daud & Mohammed, 2014). Similarly, Afolagboye et al. (2017) reported that the low hydraulic conductivity of the fine contents provided a more tortuous flow path for water to flow. All studied samples yielded hydraulic conductivities which were all in the order of 10^{-9} m/s at all different moulding water contents.

Table 6
The variation of hydraulic conductivity with steel slag content

Steel Slag Content (%)	**Hydraulic Conductivity (m/s)**
0	6.603×10^{-5}
5	1.182×10^{-6}
10	7.250×10^{-7}
15	1.952×10^{-8}

The material used in the construction of compacted liners should be strong enough to withstand shear failure caused by the load imposed by the overlying waste materials. Daniel and Wu (1993) recommended that the minimum soil strength to be used as compacted soil liners should be 200 kN/m^2. The shear strength values decreased with an increase in moulding water contents for all samples, whereas the presence of steel slag substance led to an increase in shear strength values (Figure 3).

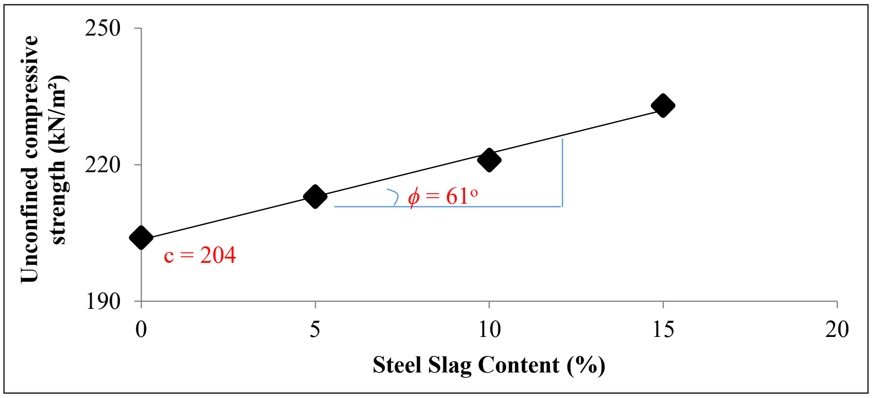


Figure 3. Variation of Unconfined Compressive Strength (UCS) with steel slag content
Shear parameters: c – cohesion & ϕ -internal friction angle

The estimated cohesion, c and the angle of internal friction, (ϕ) from the Mohr envelopes were shown in Figure 3. The increase in strength with steel slag content of up to 15% possibly attributed to the formations of cementations product through pozzolanic reactions and hydration of cementations material that coated and bound the soil particles to produce stronger matrices (Amadi et al., 2012; Nik Daud & Mohammed, 2014). Test results (Figure 3) confirmed that the soil possessed a higher strength than the recommended minimum strength when it was mixed with steel slag (Nik Daud & Mohammed, 2014).

A typical stress-strain curve for studied samples was presented in Figure 4. In this study, specimen failure was defined as the maximum axial stress or when no peak was reached during the test. Peak axial stress in the UCS test was 7.32 kN/cm^2 and it was found at 8.71% strain for 10% steel slag mixtures. Generally, the shapes of the stress-strain curves differ significantly between low and high steel slag contents.

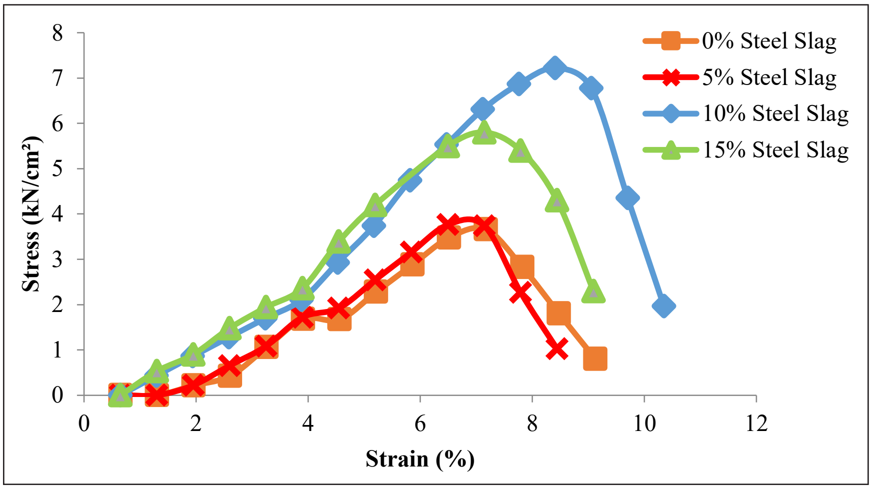


Figure 4. Relationships of stress-strain from UCS test

From the UCS test, the samples exhibit brittle characteristic in terms of rapid drop in the post-peak stress with increment in strain, which is parallel to the characteristics of the structured natural clays (Horpibulsuk et al., 2009). Figure 4 shows that only 10% steel slag are enough to fill up the voids at optimum moisture content. Thus, for steel slag mixtures, it can be seen that less than 15% mixture are able to fill up the voids of the sand. From this study, the employment of up to 10% of steel slag is appropriate for a retention pond liner. Going for higher steel slag ratios will increase the compressibility and stress of the liner material that result in light structure problems (Dafalla, 2017).

CONCLUSIONS

This study has assessed the effects of steel slag as a strength improvement admixture on mining soil in relation to the design of liners in retention pond. The soil has an adequate amount of fine particles along with good plasticity features which are vital to achieve a low hydraulic conductivity, while MDD and OMC decrease and increase, respectively, with higher steel slag contents. The low hydraulic conductivity and adequate strength, combined with their availability could have made them potential materials to be used as compacted soil liners for environmental protection. Therefore, the results have indicated that the mining soil-steel slag has promising potential as a liner material as it meets the USEPA criteria for waste-containing liner.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the financial support provided by the Ministry of Higher Education Malaysia through the Fundamental Research Grant Scheme (FRGS 5524757) and IPS 9574900 funded by Universiti Putra Malaysia.

REFERENCES

Afolagboye, L. O., Talabi, A. O., Ajisafe, Y. C., & Alabi, S. (2017). Geotechnical assessment of crushed shales from selected locations in Nigeria as materials for landfill liners. *Geotechnical and Geological Engineering, 35*(4), 1847-1858.

Amadi, A. A., Eberemu, A. O., & Osinubi, K. J. (2012). Strength consideration in the use of lateritic soil stabilized with fly ash as liners and covers in waste landfills. In D. H. Roman (Ed.), *GeoCongress 2012: State of the art and practice in geotechnical engineering* (pp. 3835-3844). Oakland, California: American Society of Civil Engineers.

ASTM D2487. (2011). *Standard practice for classification of soils for engineering purposes (Unified soil classification system).* West Conshohocken, US: Author.

Benson, C. H., Zhai, H., & Wang, X. (1994). Estimating Hydraulic Conductivity of Compacted Clay Liners. *Journal of Geotechnical Engineering, 120*(2), 366-387.

Dafalla, M. A. (2017). The compressibility and swell of mixtures for sand-clay liners. *Advances in Materials Science and Engineering, 9*(2017), 1-9.

Daniel, D. E., & Benson, C. H. (1990). Water content - Density criteria for compacted soil liners. *Journal of Geotechnical Engineering, 116*(12), 1811-1830.

Daniel, D. E., & Wu, Y. K. (1993). Compacted clay liners and covers for arid sites. *Journal of Geotechnical Engineering, 119*(2), 223-237.

Fashola M., Bgole-Jeme V., & Babalola, O. (2016). Heavy metal pollution from gold mines: Environmental effects and bacterial strategies for resistance. *International Journal Environmental Research and Public Health, 13*(11), 1047-1067.

Herrmann, I., Andreas, L., Diener, S., & Lind, L. (2010). Steel slag used in landfill cover liners: Laboratory and field tests. *Waste Management & Research, 28*(12), 1114-1121.

Horpibulsuk, S., Rachan, R., & Raksachon, Y. (2009). Role of fly ash on strength and microstructure development in blended cement stabilized silty clay. *Soils Foundation, 49*(1), 85-98.

Kusin, F. M., Rahman, M. S. A., Madzin, Z., Jusop, S., Mohamat-Yusuff, F., Ariffin, M., & Zahar, M. S. M. (2017) The occurrence and potential ecological risk assessment of bauxite mine-impacted water and sediments in Kuantan, Pahang, Malaysia. *Environmental Science & Pollution Research*, *24*(2), 1306-1321.

Kusin, F. M., Azani, N. N. M., Hasan, S. N. M. S., & Sulong, N. A. (2018) Distribution of heavy metals and metalloid in surface sediments of heavily-mined area for bauxite ore in Pengerang, Malaysia and associated risk assessment. *Catena*, *165*(2018), 454-464.

Lagos, G. I., & Geo, P. (2011). The use of bench-scale treatability studies in the design of engineered wetlands for the remediation of Acid Mine Drainage (AMD) and leachate in the vicinity of coal minesa case study in Ohio, United States. *Procedia Earth and Planetary Science*, *3*(2011), 11-16.

Li, L., Lin, C., & Zhang, Z. (2017). Utilization of shale-clay mixtures as a landfill liner material to retain heavy metals. *Materials and Design, 114*(2017), 73-82.

Masri, N, Kusin, F. M., Nik Daud, N. N., & Jamil, M. I. (2017). Properties of native soil at selinsing gold mine for geotehnical characterization of liner. In R. S. Azis & M. M. A. Kechik (Eds.), *Proceedings of the Conference on Emerging Themes in Fundamental and Applied Science* (pp. 188). Serdang, Malaysia: UPM Press.

Miguel, M. G., Barreto, R. P., & Pereira, S. Y. (2015). Analysis of aluminum, manganese, and iron adsorption for the design of a liner for retention of the acid mining drainage. *Water, Air, and Soil Pollution*, *226*(3), 1-67.

Molahid, V. L. M., Kusin, F.M., & Madzin, Z. (2018) Role of multiple substrates (spent mushroom compost, ochre, steel slag and limestone) in passive remediation of metal-containing acid mine drainage. *Environmental Technology, 40*(10), 1323-1336.

Muhamad, S. N., Kusin, F. M., Zahar, M. S. M., Mohamat-Yusuff, F., & Halimoon, N. (2017) Passive bioremediation technology incorporating lignocellulosic spent mushroom compost and limestone for metal- and sulphate-rich acid mine drainage. *Environmental Technology, 38*(16), 2003-2012.

Nik Daud, N. N., & Mohammed, A., S. (2014). Strength assessment in the use of granite residual soil treated with palm oil fuel ash (POFA) as hydraulic barrier in landfills. In *7th International Congress on Environmental Geotechnics* (p. 1485). Barton, Canberra: Engineers Australia.

Qian, X., Koerner, R. M., & Gray, D. H. (2002). *Geotechnical aspects of landfill design and construction.* Upper Saddle River, New Jersey: Prentice Hall.

Rauch, A., Harmon, J., Katz, L., & Liljestrand, H. (2002). Measured effects of liquid soil stabilizers on engineering properties of clay. *Transportation Research Record, 1787*(1), 33-41.

Road Engineering Standards Policy Committee. (1990). *British standard methods of test for soils for civil engineering purposes*. London, United Kingdom: Author.

Rowe, R. K., Quigley, R. M., Brachman, R. W., & Booker, J. R. (2004). *Barrier systems for waste disposal facilities.* Florida, US: CRC Press.

United States Environmental Protection Agency. (1989). *Requirements for hazardous waste landfill design, construction and closure*. Cincinnati, Ohio: Environmental Protection Agency.

Wagner, J. F., & Schnatmeyer, C. (2002). Test field study of different cover sealing systems for industrial dumps and polluted sites. *Applied Clay Science, 21*(1-2), 99-116.

Zahar, M. S. M., Kusin, F. M., & Muhammad, S. J. (2015). Manganese removal from aqueous solution by steel slag: Kinetic and equilibrium studies. *International Journal of Scientific Research in Science, Engineering and Technology, 5*(1), 20-29.

Pertanika J. Sci. & Technol. 27 (4): 2317 - 2331 (2019)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Distribution of Dissolved and Particulate Fe in an Estuarine System at Bagan Pasir, Perak

Syaza Syafiqah Muhamad Kamel[1], Khairul Nizam Mohamed[1,2]*, Ferdius@ Ferdaus Mohamat Yusuff[1] and Abd Muhaimin Amiruddin[3]

[1]*Department of Environmental Sciences, Faculty of Environmental Studies, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia*
[2]*Environmental Forensic Research Unit (ENFORCE), Faculty of Environmental Studies,Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia*
[3]*Department of Environmental Management, Faculty of Environmental Studies, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia*

ABSTRACT

The distribution of dissolved iron (dFe) and particulate iron in the estuarine system was studied where *in-situ* water sampling stations were selected at Bagan Pasir, Perak. The concentration of dFe was 1.17±0.28 mg/L on average at the estuary, while in freshwater samples its concentration was 0.08±0.00 mg/L. This study found that the concentration of particulate Fe in freshwater system was higher than that in the estuary system. The Fe concentration was 0.95±0.03 mg/kg and 0.80±0.18 mg/kg at the freshwater and estuary systems, respectively. We have applied a distribution coefficient (K_D) in order to quantify the partitioning of Fe between the particulate (>0.45 µm) and dissolved (<0.45 µm) phases. The K_D value of Fe in the freshwater system was much higher compared to that in the estuary, where the K_D value was 12.18 in freshwater and 0.48 -1.31 in the estuary. This indicated that Fe in this area has a strong affinity with the particulate phase in a freshwater system and this is probably due to the anthropogenic input. On the other hand, Fe may exist in colloidal or organic ligands which probably originate from surface sediment or a biological process in water columns at the estuarine area.

Keywords: Dissolved iron, estuary, particulate iron, seawater, voltammetry

ARTICLE INFO

Article history:
Received: 03 December 2018
Accepted: 30 January 2019
Published: 21 October 2019

E-mail addresses:
syaza.syafiqah.muhamad@gmail.com (Syaza Syafiqah Muhamad Kamel)
k_nizam@upm.edu.my (Khairul Nizam Mohamed)
ferdius@upm.edu.my (Ferdius@Ferdaus Mohamat Yusuff)
muhai@upm.edu.my (Abd Muhaimin Amiruddin)
* Corresponding author

INTRODUCTION

Iron (Fe) acts as an important element as well as a micronutrient for marine organisms and has a key role in sustaining estuary biogeochemistry (Gledhill et al., 2004). The

ISSN: 0128-7680
e-ISSN: 2231-8526

bioavailability and biogeochemistry of Fe depends on the probability to form either in metal complexes or colloids (Lohan et al., 2005). Therefore, the concentration of iron influences the level of atmospheric CO_2 uptake via phytoplankton growth to maintaining the marine ecosystem. In oxygenated and highly saline waters iron has low solubility; thus, this is the reason why Fe concentration is low. Otherwise, in estuaries and rivers, Fe is mainly present in association with dissolved organic matter in areas close to the ocean. A previous research (Asmala et al., 2013) had studied characterization of riverine dissolved organic matter and the significant role it played in Fe transport capacity of the river waters and extent to which riverine iron might be maintained in solution in the marine environment. It was suggested that an increasing amount of dFe could enhance the productivity of marine systems.

Therefore, some of the dissolved species may react with suspended particles. Fe predominantly occurs in the particulate form (Costa et al., 2016; Gledhill et al., 2004) with lower dissolved concentration. Anthropogenic and many other human activities produce residue and toxic substances (mostly heavy metals) that are eventually discharged into rivers and estuaries. These subsequently flow to the open ocean (Syvitski et al., 2005). Thus, these metals are present in different components of an aquatic system such as: water body; suspended matter; bed sediments; and biota. In this way, they remain in the medium as highly toxic micro-pollutants and cause adverse effects (Schnoor, 1996). In order to better understand what controls estuarine primary production, we need to understand clearly the distribution and fractionation of trace metals in seawater. Fe is reported to be rendered more complex by ligands that can relate to some specific chemical compounds in coastal waters (Avendaño et al., 2016; Gledhill et al., 2004; Huber et al., 2011; Mawji et al., 2011; Velasquez et al., 2011).

However, most of the current trace metal studies performed in Malaysia were carried out in sediment and freshwater (Ahmad et al., 2009; Adiana et al., 2014; Koh et al., 2015; Suratman et al., 2009; Yusof et al., 2009; Yap et al., 2011). There were also a few studies conducted of Fe in estuarine areas. Billah et al. (2016) studied the distribution (Fe, Zn, Cu and Mn) and water characteristics at Miri estuary, Sarawak, by using Atomic Absorption Spectrophotometer (AAS). An earlier study by Mohamed et al. (2016) focused on fractioning of Mn, Fe and Cu in seawater at Port Dickson, Negeri Sembilan by applying Chelex 100 to a preconcentrate before determination by Inductive Couple Plasma – Mass Spectrometry (ICP-MS). However, most of these methods require complex and expensive analytical equipment, thereby limiting their potential application. This could affect the limitation of understanding of biogeochemistry of dissolved Fe in an estuarine system. Therefore, a lack of dissolved Fe data might be due to the use of an analytical method where the sample itself (seawater) has a high metric and contains higher levels of salt. There are challenges involved in analysing the salty water sample directly. It is vital to have more knowledge of estuarine systems since most estuarine areas are the main economic sources of fresh seafood stock. This stock will be supplied to markets where humans will be the final consumer. This study was proposed so as to determine dissolved Fe in estuary water by using the current developed method (electrochemistry method), along with determination of particulate Fe. Due to the high salt content, this technique can be undertaken directly in seawater. The pre-concentration procedure is performed and followed by a potential scan towards more negative potentials and the cathodic (reductive) currents are then determined. This method is based on the measurement of current response as a function of the potential applied to a voltammetric cell.

The objectives of this study were to identify the distribution of dissolved Fe (dFe) and particulate Fe (Fe_p) in an estuarine system at Bagan Pasir, Perak and also to determine the influence of *in-situ* parameters on its fractionation distribution at an estuarine location. This is important in relation to updating our knowledge and increasing our understanding of the effect of anthropogenic and natural activities on the Fe biogeochemistry cycle and allowing us to monitor the productivity of our estuarine systems.

MATERIALS AND METHODS

Acid Wash

Sample container (1L, low-density polyethylene (LDPE)) bottles were cleaned by soaking in 30 % HNO_3 overnight followed by thorough rinsing with Milli Q to remove all the metal contaminants. All sample bottles used in this study were cleaned according to a standard protocol (Achterberg et al., 2001).

Sampling Area

The Bagan Pasir estuary is located west of Teluk Intan (Figure 1), in the state of Perak, Malaysia. The majority of the fresh water entering the estuary originates from the Tanjung Karang (Perak's main river) which drains a catchment area of approximately 14900 km^2.

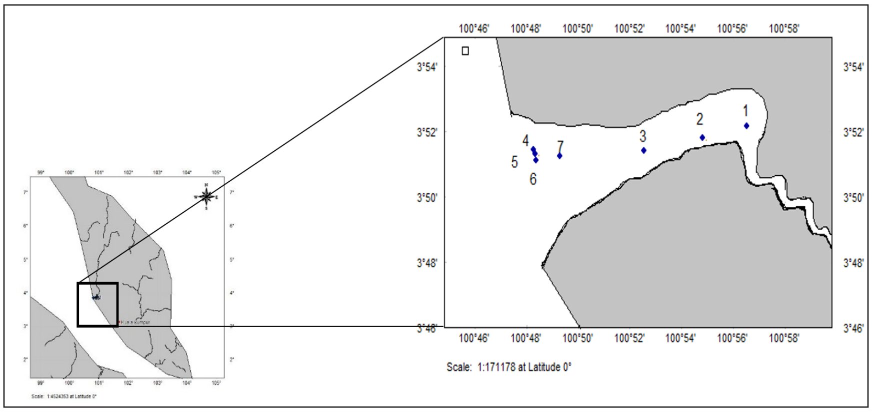


Figure 1. The sampling location at Bagan Pasir estuary, Perak on January 2018

The Bagan Pasir area is well known as a fishing village, producing about 30 different seafood supplies including pomfret, prawn, shrimp, and cockles. Most of the suppliers at the market around Selangor visit this place in order to refresh their stocks (Kelly, 2016).

Sampling Activity

Sampling was carried out in January 2018 at Bagan Pasir, Perak. The water samples were collected from seven stations (Figure 1). Our sampling stations were divided into three (3) zones based on salinity, i.e.: Zone 1 (St.1, salinity of 7.52 ppt); Zone 2 (St.3 and St.2, salinity of 13.18 ± 0.05 ppt); and Zone 3 (St.7, St.6, St.5 and St.4 salinity of 16.42 ± 2.25 ppt) (Figure 1). This classification was applied in order to identify the possible influences of salinity on Fe distribution in the water column as suggested by the United States Environmental Protection Agency (US EPA, 2006). Stations 1, 2 and 3 were a distance of 1km apart from each station; while the other stations were located near to the estuaries area with a distance of 500m apart from each station. A 1 L sample was collected at middle depth at each station by using Van Dorn water sampler and was then dispersed into a PTFE bottle for further analysis. The reading of *in-situ* parameters was recorded by using, specifically: pH meter Thermo Orion (Model 230); Dissolved Oxygen (DO) meter (YSI Model 52); salinity meter (YSI 30); Conductivity meter (Thermo Scientific Orion 3-star Plus); and Turbidity meter (Thermo Orion Model AQ4500). The in-situ parameters used during our sampling are as shown in Table 1 following.

Table 1
The in-situ parameters data during sampling at Bagan Pasir, Perak

Station	Coordinate	Sal. (ppt)	Cond. (mS)	Turbidity (NTU)	Temp. (^oC)	pH	DO (mg/L)	dFe (mg/L)	Fe_p (mg/kg)	K_D
1	3.8696 100.9420	7.52	14905.00	127.00	28.80	7.05	2.81	0.08±0.04	0.95±0.03	12.18
2	3.8635 100.9136	13.21	23806.00	155.00	29.10	6.60	3.32	1.15±0.62	0.55±0.03	0.48
3	3.8567 100.8755	13.14	30794.00	97.00	29.30	6.80	6.03	0.70±0.02	0.91±0.02	1.31
4	3.8576 100.8046	13.20	39855.00	75.00	29.00	7.27	4.51	1.26±0.60	0.87±0.00	0.69
5	3.8553 100.8054	17.43	24230.00	58.00	31.60	7.45	5.26	1.26±0.64	0.71±0.02	0.56
6	3.8522 100.8060	18.41	33337.00	23.00	29.40	7.31	5.31	1.55±0.67	1.06±0.04	0.69
7	3.8542 100.8215	16.87	38261.00	0.00	29.90	7.39	4.14	1.12±0.55	0.72±0.01	0.64

Sample Treatment

About one litre of water samples (1L) was sampled using filter vacuum pump as described by Bruland et al. (2003) and Achterberg et al. (2001). A 500ml of sample for total dissolved analysis was acidified to pH 1.99 to 2.10 (a final concentration of 10%) using pure 65 % HNO_3 (RCI Premium). Total suspended solid (TSS) samples obtained from filtration of a 1L sample were allowed to dry for several days until a constant weight was attained.

Particulate Fe Analysis

A process of acid digestion was suggested by Adiana et al. (2014) with a modification. In this modification, a digestion of TSS samples was conducted by using 2 ml of HF, 3 ml of HCl and 4 ml of HNO_3 in a closed Teflon vessel on a hotplate for 30 min duration. This was followed by rinsing of the sample using 2% HNO_3 by open digestion process and adjusting the volume until 10 ml. It was then heated up again for approximately 2-3 min so as to obtain a good mixture of solution and finally made up to 25 ml with Milli-Q water. The digested sample was kept in centrifuge tubes at 4°C before determining particulate Fe by using Induced Coupled Plasma Mass Spectrometry (ICP-MS).

Inorganic Dissolved Fe Analysis

All sample preparation was carried out in a Class 100 laminar airflow bench at room temperature. Milli-Q (Millipore U.K) of 18.2MΩ·cm resistivity was used to prepare reagents and dilutions. Iron solutions of different concentrations were prepared by diluting 1000 ppm atomic absorption standard solutions (BDH, UK) which were then acidified to pH 2 with HCl. HCl (Merck) and ammonia (EMSURE, Merck) were purchased from Fisher Scientific and were used to adjust the pH. A stock solution of 0.02 M 2-(2-Thiazlylazo)-p-cresol (TAC) was prepared by dissolving TAC in 100% Methanol. The TAC was obtained from Sigma-Aldrich and had a purity rating of 97%. The pH buffer contained 1 M boric acid in 0.3 M ammonia in Milli-Q and gave pH 8.05-8.15.

Inorganic dissolved Fe analysis was conducted by using the method devised by van der Berg and Huang (1984) with some modification. A sub-sample (15 ml) of treated water sample (pH: 2) was pipetted into a polarographic cell and the pH was adjusted to 8.0 by ammonia hydroxide solution. Following this, there was an addition of 1µM (100µl) of 1M borate buffer. The solution was deaerated by bubbling with pure nitrogen gas (99.99%), and 150µl of 1mM of TAC solution were then added. A new mercury drop was extruded and the potentiostat was then switched on and set to 0.1V. The stirrer was started in order to initiate the collection which was carried out for 8 min. A negative scan was then made 10 s after stopping the stirrer using a fast linear sweep of 50 Mv/s. Deposition time was about 1 min in duration. The smooth reduction peak of Fe lies between -0.35 V to -0.4 V.

RESULTS AND DISCUSSIONS

The results of the Fe concentrations are presented in Table 1. We found that the salinity at St. 1 was 7.52 ppt (Table 1) which was the lowest result. St. 2, 3 and 4 were ranged between 13.14 to 13.21 ppt (Table 1) and St. 5, 6 and 7 recorded the highest salinity levels, namely, 16.87 to 18.41 ppt (Table 1). Present collected data had shown that the salinity increased from St. 1 through to St. 7 (Figure 2). Salinity refers to the effect on a biogeochemical process through chemical and physical aspects. This occurs through salinity where salt within seawater interacted with iron, thereby having the potential to experience ion exchange (Bai et al., 2012). Findings from a volunteer estuary monitoring manual stated that salinity in an estuary played a role involving flocculation of particles, where the particles that entered an estuary dissolved in fresh river water and collided with salt water. Hence, they may flocculate resulting in increased turbidity thereby influencing the distribution of Fe. From this, our results found that the dFe increase with salinity increase has been proven from the study of Jilbert et al. (2018) at Boreal estuary.

A pH at St.1 was found at 7.05 (Table 1) followed by St. 2, 3 and 4 at 6.60 to 7.27 (Table 2) which is slightly acidic. At St.7, the pH was 7.31 to 7.45 (Table 1) which is slightly more alkaline to the open sea area. This occurs because most of the dissolved minerals

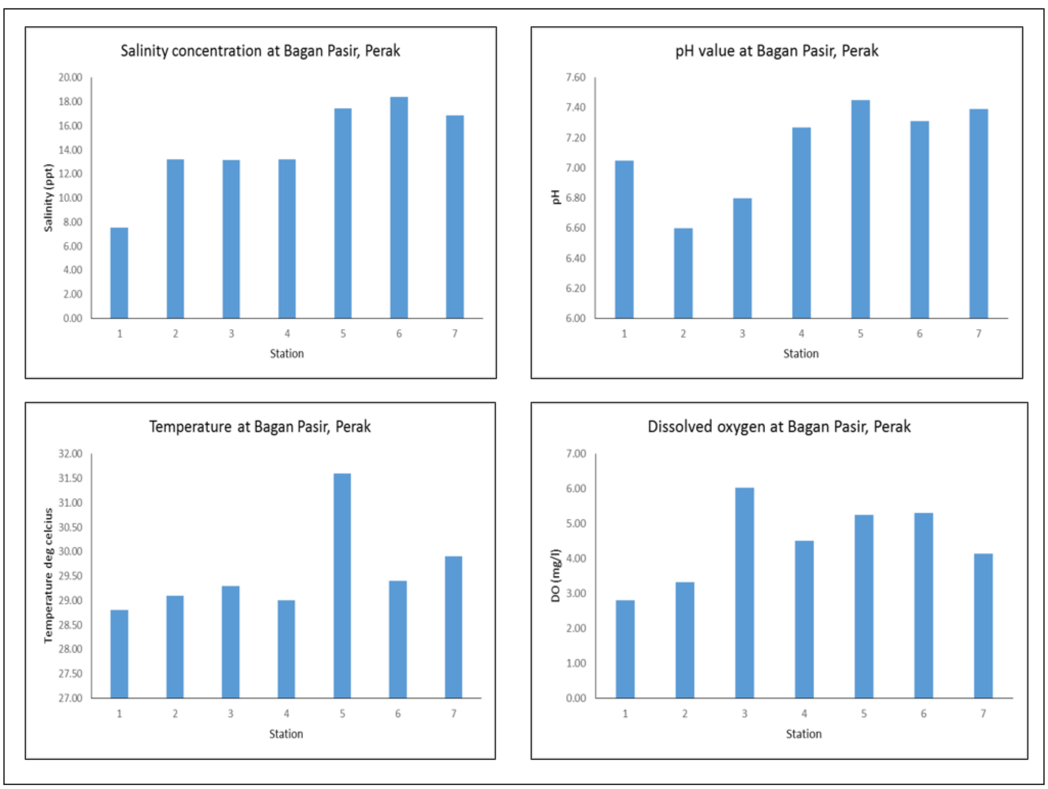


Figure 2. The in-situ parameter pattern and distribution at sampling station at Bagan Pasir, Perak

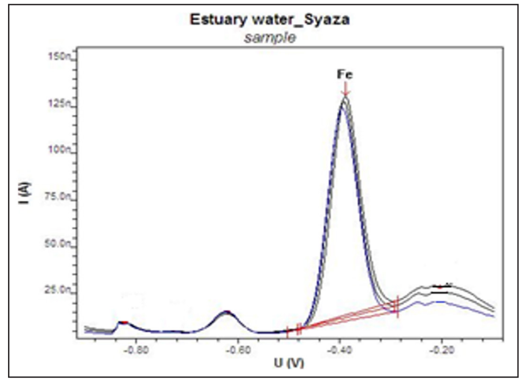
(dissolve CO_2 turns to HCO_3) are picked up by groundwater and carried eventually to the ocean (Sophie, 2017). The alkalinity of the water body indicated the presence of CO_3^{2-}, Ca^{2+} and Mg^{2+}. The acidity may cause a photosynthetic process and decomposition of organic matter mostly from urban runoff (Billah et al., 2016). During our sampling, the temperature recorded was also the same; ranging between 28.80 to 31.60℃. The temperate zone might not influence other parameters; while some research reported that the influence of temperature on metal transportation was not evident, but could be influenced by salinity changes (Aston et al., 2010; Biesuz et al., 1998; Zhang et al., 2013). Mokhtar et al. (2009) stated that change of temperature on surface water was a function of heat exchange on the earth's surface; hence, our recorded temperature was probably influenced by the tidal cycles since our sampling was carried out at high flow cycles.

The concentration of dissolved oxygen (DO) at St.1 was 2.81 mg/L, St.2 , 3 and 4 were 3.32 to 6.03 mg/L and 4.14 to 5.36 mg/L at St.5, 6 and 7 (Table 1), respectively. This data shows an increase of DO from St.1 to St.7 (Figure 2). This could be the reason for a slower oxidative precipitation rate and scavenging of dissolved heavy metal ions by iron hydroxide formed from Fe cations simultaneously released from pore waters under low DO (Atkinson et al, 2007; Santana-casiano et al, 2004). Whereas at high DO, oxygen penetration was much stronger, Fe oxidative precipitation generated hydroxide or oxide would adsorb more dissolved heavy metal, thus leading to reduced metal release (Zhang et al., 2017).

The dFe concentration in the sample was determined by standard additions to 1ppm of Fe (III), which gave a peak current of about 114.58 nA (Figure 3). Measurements were made with 15 ml of samples which were buffered to pH 8.05 by 1 M boric acid, in the presence of 10µM TAC as an added ligand to be binding to the natural Fe in sample. Collection was carried out using HMDE for 8 min fraction of the Fe-TAC which was adsorbed on the surface of the Hg drop. A potential scan direction is aimed toward more negative potentials and the resulting current is measured at -0.1 V. This was followed by using Differential Pulse modulation at a pulse rate of 10 pulse/s and scan rate at 10Mv/s. It was found that the maximum peak height which could produce at high iron concentrations was about 118.20 nA (Figure 3(b)). Where the resultant current-potential stripping voltammogram provides quantitive information, the height of the peak is of the standard addition method and qualitative information. The potential of the peak is an indication of the Fe analyzed. This technique is sufficiently sensitive for the direct determination of dissolved iron in seawater as it has a concentration in oceanic regions in nM range. Hence, voltammetric analysis (Figure 3(a)) may vary between samples of different ionic strength and containing different concentrations of surfactants and natural trace metal complexing organic ligands.

The dFe concentration at St.1 was 0.08±0.04; St.2, 3 and 4 were ranged between 0.70±0.02 to 1.26±0.62 mg/L and St. 5, 6 and 7 were 1.12±0.55 to 1.55±0.67 mg/L (Table 2), respectively, in our study area. Most of the stations which were located at the estuary

(a)



(b)

Figure 3. (a) 797 Voltammetric Analysis Computrace Stand, (b) determination of dissolved Fe using DPCSV, 10 pulse/s, 10 Mv/s. The collection time was 1 min.

had higher levels of dFe concentration (Figure 4). This indicated that the concentration of dFe was increased by salinity and DO. This could suggest that Fe input was occurring naturally in the environment. Further, it could be mobilized by anthropogenic activities consisting of mining and industrial effluents discharging from rivers into natural ecosystems like forest, mangroves, estuaries and ocean (Larison et al., 2000). Our dFe in mean was (1.0±0.2) $\times 10^3$ µg/L in an estuary, which is lower than that of other studies (Table 3). This was a comparison from previous studies that were mostly focused in river areas, which are a direct source of heavy metal.

Table 2
Range of in-situ parameters and concentration of Fe in dissolved and particulate phase at Bagan Pasir, Perak

Station	Salinity (ppt)	Temperature (^{o}C)	pH	DO (mg/L)	dFe (mg/L)	Fe_p (mg/kg)	K_D
1	7.52	28.80	7.05	2.81	0.08±0.04	0.95±0.02	12.18
2, 3 and 4	13.14-13.21	29.00-29.30	6.60-7.27	3.32-6.03	0.70±0.02-1.26±0.62	0.55±0.03-0.91±0.02	0.48-1.31
5, 6 and 7	16.87-18.41	29.40-31.60	7.31-7.45	4.14-5.36	1.12±0.55-1.55±0.67	0.71±0.02-1.06±0.04	0.56-0.69

Results from previous studies indicated that the metal concentrations were comparable. There was also general acceptance on the area that had been polluted by agricultural activities, as well as industrial, effluent discharge and human activities (Table 3). However, the fraction of riverine Fe present as discrete ferrihydrite particles had been shown to increase with rising pH, as was typically observed with increasing stream order towards the coastal zone (Neubauer et al., 2013). It was found that Fe probably did exist in colloidal form or being bound with organic ligands.

Table 3
Average concentration of Fe in dissolved (ug/l) and particulate phase (ug/l) of river water samples as noted in previous studies

Rivers	Dissolved Fe (dFe)	Particulate Fe (Fe_p)	**References**
Setiu	$(2.0\pm3.4)\times10^2$	10±13	Koh, 2015
Langat	$(1.3\pm2.3)\times10^3$	Nd	Yusof et al., 2009
Kelantan	Nd	$(3860\pm30)\times10^3$	Ahmad, 2009
World Average	50	5.8±4.8	Viers et al, 2009; Meybeck et al., 1996
Bagan Pasir estuary	$(1.0\pm0.2)\times10^3$	$(0.80\pm0.18)\times10^3$	This study

Note: *nd = no data available

Based on the correlation of Fe in a dissolved phase with in-situ parameters, we found that dFe had a significant linear relationship with salinity. In estuarine environments, elevated electrolyte strength along salinity gradients induces the flocculation of DOM (Sholkovitz et al., 1978) and Fe (Boyle et al., 1977) from river waters. This phenomenon is usually explained by the cation-induced aggregation of colloidal humic substances, which carry a net negative surface charge (Eckert & Sholkovitz, 1976). This suggests a destabilization of colloidal iron as ionic strength increases (Byrd et al., 1990). In contrast, we found that dFe did not have a significant relationship between pH, DO and temperature after $P\geq0.05$. In the case of Fe_p, there is an indication that all the in-situ parameters do not have any influence upon Fe_p. This could indicate that there is little or no remobilization of iron in the sediments in this area.

Table 4
Correlation coefficient between dissolved and particulate Fe concentration with in-situ parameter at Bagan Pasir Perak

Parameter	**R value (dFe)**	**R value (Fe_p)**
Salinity	0.88	0.09
pH	0.37	0.30
Temperature	0.37	0.33
DO	0.46	0.32

Concentration of Fe in particulate form at the Bagan Pasir estuary can be seen in Table 1. Fe_p concentration ranged from 0.95±0.03 to 1.06±0.04 mg/kg in inverse relation with river levels (estuary to river). The results showed a slight variation between each station. High concentrations of particulate Fe were found at the estuary area (St.1); this is due to the turbidity maximum area and related to the estuarine dynamic circulation (Barak & Nurit, 1997). It is represented by a simple model of seawater transport and settling of particulate metal from the surface to the deep layer of water and can be carried back to the upstream and remixed with surface water (Postma, 1967). Both instances of Fe "inputs" concentration were below the values obtained in the river in all stations, which suggests that this metal comes from upstream. It is probable that during the last two years, the stream was flooded by river waters in November, 2016 (Kelly, 2016).

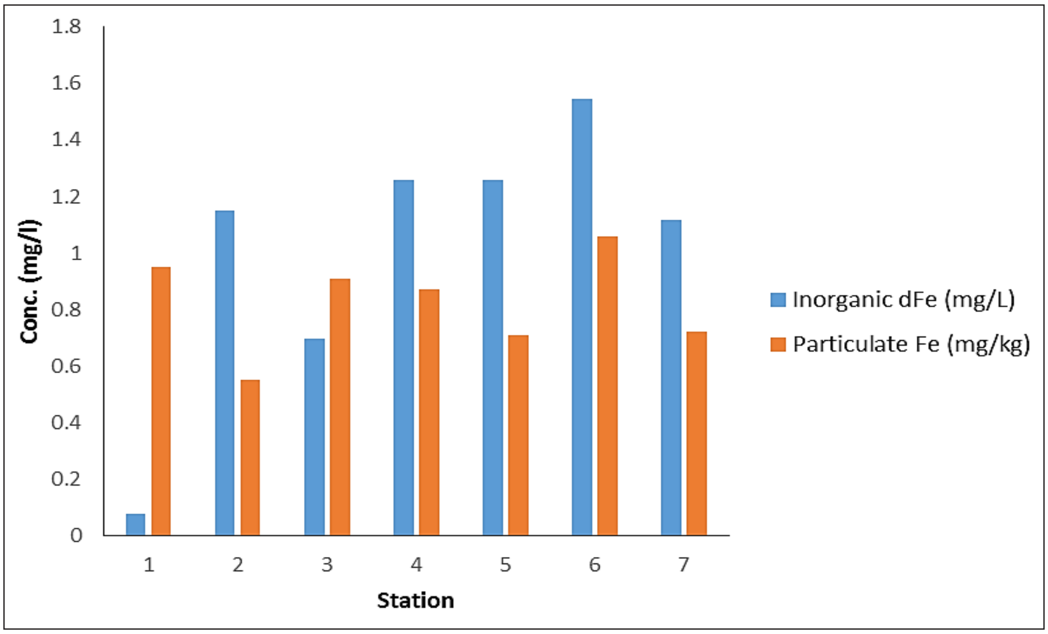


Figure 4. Distribution of dissolved and particulate Fe at Bagan Pasir estuary

A pattern of Fe was observed for this distribution in both dissolved and particulate phase as shown in Figure 4. Where the particulate phase showed a higher concentration than the dissolved phase of Fe ($Fe_p \geq dFe$) at St.1 and St.3, at other stations it displayed $dFe \geq Fe_p$ (Fig. 5). In order to find out the partitioning of a metal between the particulate (>0.45 µm) and dissolved (<0.45 µm) phases, we applied a distribution coefficient, K_D (Benoit et al., 1994; Benoit & Rozan, 1999; Munksgaard & Parry, 2001; Nguyen et al., 2005):

$$K_D = \frac{[\,particulate\ metal\ concentration](ug/kg)}{[\,dissolved\ metal\ concentration](ug/l)}$$

K_D is a measure of the tendency of an element to be associated and transported with the particulate phase. High particle reactivity for a metal would tend to increase that metal's K_D value. A portioning coefficient in this study shows that St.1 is 12.18, St.2, 3 and 4 are 0.48 to 1.31 and St.5, 6 and 7 are 0.56 to 0.69 (Table 1), respectively. It indicated that St. 2 to St. 7 and zone 3 were lower than St.1. This could suggest that Fe at St.1 might have strong affinities for TSS and also might represent the competition between dissolved and particulate metals binding sites. This is due to anthropogenic activities and urban runoff directly from a river. However, St.2 to St.7 showed transportation particulate to the dissolved phase. This is due to Fe in the particulate phase re-suspending from anoxic zones and coming into contact with oxygenated water where sediment suspension occurred (Hou et al, 2013). The particulate iron is dissolved by bacteria, which also release ligands for iron ions complexation and keeps them in the solution. The oxidation state of iron in an aquatic system and its redox processes are determined by the dissolved oxygen concentration. Redox reactions, complexation with inorganic and organic ligands, adsorption and precipitation are involved in the chemistry of iron in natural waters; whereas at high DO, oxygen penetration was stronger, and Fe and Mn oxidative precipitation was more rapid. Hence, the generated hydroxide or oxide would adsorb more dissolved heavy metals, thus leading to reduced metals release. Therefore, the result of both zones having low K_D values was probably due to most of the Fe possibly existing in organic ligands or colloidal forms (Byrd et al., 1990).

The organic ligands were thought to stabilize the metals through struggling with binding sites on settling particles (Mohamed et al., 2011). These results were in line with the kinetic iron studies carried out by Mayer (1982) which showed that iron agglomerates in the form of colloids during mixing of fresh water and saline water while following two (2) stages of a kinetic process. There is a first, fast stage (zone 1) with first order kinetic, relating to the interaction of iron with high molecular mass of organic matter followed by a second, slower stage (zones 2 and 3). This stage has a second order kinetic corresponding to the formation of colloidal particles which would depend on temperature and turbidity (quantity of suspended particles). This could also relate with our turbidity data, where the value decreased from a river (St.1) to 127.00 NTU. The lowest value along the estuarine area (St. 2, 3, 4, 5, 6 and 7) is 23.00 NTU respectively (Table 1).

For instance, the high value of K_D at St.1 might originate from anthropogenic activities directly linked to the river possibly resulting in multiple sources of organic matter. This indicates that most of the Fe exists in particulate phase rather than dissolved phase at St.1. Moreover, the characteristic of freshwater could reveal the condition of surrounding water. Many researchers reported that at low pH (6.0 or less), the metal release was more obvious than at middle (7.0) or high pH (more than 7.0); however, low pH value instances are very rare in a natural environment (Butler, 2009; Jing et al., 2007; Pérez-Esteban et al., 2013; Watmough et al., 2007; Yang et al., 2006). This provides an explanation to our data at Zone 1 (fresh water).

Furthermore, domestic sewage as one of the input contributors of anthropogenic activities could be the main factor. Residents in the Bagan Pasir area stated that they had no proper rubbish disposal system (Kelly, 2016). Besides, dFe at zone 1 was found to be lower, probably due to the presence of low dissolved oxygen (DO); hence, the oxidative precipitation rate would become slower. Further, scavenging of dissolved Fe ions form iron hydroxide by Fe cations released from pore water as they diffuse from sediment to water (Atkinson et al., 2007; Santana-Casiano et al., 2004).

CONCLUSIONS

Dissolved Fe and particulate Fe samples along some stations at an estuarine system at Bagan Pasir were observed in this study. Our K_D value indicated a decrease along the station due to most of the metal being predominantly in dissolved fraction with the low value of particulate fraction. Our data found that concentration of dFe was lower than Fe particulate in fresh water as opposed to the estuary area. This could suggest the possibility of existence of Fe in an estuary mostly in colloidal or in organic ligand forms. However, further investigation on these speciation of dissolved Fe is needed in order to improve our knowledge on iron biogeochemistry cycle in the estuarine system.

ACKNOWLEDGEMENTS

The authors would like to thank all the crew and laboratory staff during sampling and analysis for this study. We also express our sincere gratitude to the UPM grant (GP/2018/9608500), (FRGS/1/2017/WAB05/UPM/02/1) and fund by SGS/GRF/GS50644. Ultimately, sincere thanks are extended to the Faculty of Environmental Science, Universiti Putra Malaysia.

REFERENCES

Achterberg, E. P., Holland, T. W., Bowie, A. R., Mantoura, R. F. C., & Worsfold, P. J. (2001). Determination of iron in seawater. *Analytica Chimica Acta, 442*(1), 1-14.

Adiana, G., Shazili, N. A. M., Marinah, M. A., & Bidai, J. (2014). Effects of northeast on trace metal distribution in the South China Sea off Peninsular Malaysia. *Environmental Monitoring and Asssessment, 186*(1), 421-431.

Ahmad, A. K., Musrifah, I., & Othman, M. S. (2009). Water quality and heavy metal concentrations in sediment of Sungai Kelantan, Kelantan, Malaysia: A baseline study. *Sains Malaysiana, 38*(4), 435-442.

Asmala, E., Autio R., Kaartokallio, H., Pitkänen, L., Stedmon, C. A., & Thomas, D. N. (2013). Bioavailability of riverine dissolved organic matter in three Baltic Sea estuaries and the effect of catchment land-use. *Biogeosciences, 10*(11), 6969-6986.

Aston, J. E., Apel, W. A., Lee, B. D., & Peyton, B. M. (2010). Effects of cell condition, pH, and temperature on lead, zinc, and copper sorption to Acidithiobacillus caldus strain BC13. *Journal of Hazardous Materials, 184*(1-3), 34-41.

Atkinson, C. A., Jolley, D. F., & Simpson, S. L. (2007). Effect of overlying water pH, dissolved oxygen, salinity and sediment disturbances on metal release and sequestration from metal contaminated marine sediments. *Chemosphere, 69*(9), 1428-1437.

Avendaño, L., Gledhill, M., Achterberg, E, P., & Rerolle, V. (2016). Influence of ocean acidification on the organic complexation of iron and copper in Northwest European shelf seas: A combined observational and model study. *Frontiers in Marine Science, 3*(2016), 58-77.

Bai, J., Xiao, R., Zhang, K., & Gao, H. (2012). Arsenic and heavy metal pollution in wetland soils from tidal freshwater and salt marshes before and after the flow-sediment regulation regime in the Yellow River Delta, China. *Journal of Hydrology, 450*(2012), 244-253.

Benoit, G., & Rozan, T. F. (1999). The influence of size distribution on the particle concentration effect and trace metal partitioning in rivers. *Geochimica et Cosmochimica Acta, 63*(1), 113-127.

Benoit, G., Oktay-Marshall, S. D., Cantu, A., II, Hood, E. M., Coleman, C. H., Corapcioglu, M. O., & Santschi, P. H. (1994). Partitioning of Cu, Pb, Ag, Zn, Fe, Al and Mn between filter-retained particles, colloids and solution in six Texas estuaries. *Marine Chemistry, 45*(4), 307-336.

Biesuz, R., Pesavento, M., Gonzalo, A., &Valiente, M. (1998). Sorption of proton and heavy metal ions on a macroporous chelating resin with an iminodiacetate active group as a function of temperature. *Talanta, 47*(1), 127-136.

Billah, M. M., Kamal, A. H. M., Hoque, M. M., & Bhuiyan, M. K. A. (2016). Temporal distribution of water characteristics in the Miri estuary, east Malaysia. *Zoology and Ecology, 26*(2), 134-140.

Bruland, K. W., & Lohan, M. C. (2003). Controls of trace metals in seawater. *The Oceans and Marine Geochemistry, 6*(2003), 23-47.

Butler, B. A. (2009). Effect of pH, ionic strength, dissolved organic carbon, time, and particle size on metals release from mine drainage impacted streambed sediments. *Water Research, 43*(5), 1392-1402.

Byrd, J. T., Lee, K. W., Lee, D. S., Smith, R. G., & Windom, H. L. (1990). The behavior of trace metals in the Geum Estuary, Korea. *Estiaries, 13*(1), 8-13.

Costa, K. M., McManus, K. F., Anderson, R. F., & Ren. H. (2016). No iron fertilization in the equatorial Pacific Ocean during the last ice age. *Nature, 529*(7587), 519-522.

Gledhill, M., Mccormack, P., Ussher, S., Achterberg, E. P., Mantoura, R. F. C., & Worsfold, P. J. (2004). Production of siderophore type chelates by mixed bacterioplankton populations in nutrient enriched seawater incubations. *Marine Chemistry, 88*(1-2), 75-83.

Herut, B., & Kress, N. (1997). Particulate metals contamination in the Kishon River Estuary, Israel. Marine *Pollution Bulletin, 34*(9), 706-711.

Hou, D., He, J., Lü, C., Sun, Y., Zhang, F., & Otgonbayar, K. (2013). Effects of environmental factors on nutrients release at sediment-water interface and assessment of trophic status for a typical shallow lake, Northwest China. *The Scientific World Journal, 2013*, 1-16.

Huber, S. A., Balz, A., Abert, M., & Pronk, W, (2011). Characterisation of aquatic humic and non-humic matter with size-exclusion chromatography-organic carbon detection-organic nitrogen detection (LC-OCD-OND). *Water Research, 45*(2), 879-885.

Jilbert, T.., Asmala, E., Schröder, C., Tiihonen, R., Myllykangas, J. P., Virtasalo, J. J., … Hietanen, S. (2018). Impacts of flocculation on the distribution and diagenesis of iron in boreal estuarine sediments. *Biogeosciences, 15*(2018), 1243-1271.

Jing, Y. D., He, Z. L., & Yang, X. E. (2007). Effects of pH, organic acids, and competitive cations on mercury desorption in soils. *Chemosphere, 69*(10), 1662-166.

Kelly, T. (2016). A little place by the sea. Retrieved Apr 26, 2016, from https://www.thestar.com.my/metro/focus/2016/04/26/a-little-place-by-the-sea-the-tightknit-community-of-the-bagan-pasir-fishing-village-has-much-to-off/

Koh, M. K., Suratman, S., & Tahir, N. M. (2015). Dissolved and suspended particulate metals in Setiu River Basin, Terengganu. *Sains Malaysiana, 44*(7), 957-964.

Larison, J. R., Likens, G. E., Fitzpatrick, J. W., & Crock, J. G. (2000). Cadmium toxicity among wildlife in the Colorado rocky mountains. *Nature, 406*(2000), 181-183.

Lohan, M. C., Aguilar-Islas, A. M., Franks, R. P., & Bruland, K. W. (2005). Determination of iron and copper in seawater at pH 1.7 with a new commercially available chelating resin, NTA superflow. *Analytica Chimica Acta, 530*(1), 121-129.

Mawji, M., Gledhill, M., Milton, J. A., Zubkov, M. V., Thompson, A., Wolff, G. A., & Achterberg, E. P. (2011). Production of siderophore type chelates in Atlantic Ocean waters enriched with different carbon and nitrogen sources. *Marine Chemistry, 124*(1-4), 90-99.

Meybeck, M., Friedrich, G., Thomas, R., & Chapman, D. (1996). Rivers. In D. Chapman (Ed.), *Water quality assessments - A guide to use biota, sediments and water in environmental monitoring* (pp. 246-324). Cambridge, Great Britain: University Press.

Mohamed, K. N., Steingenberger, S., Nielsdottir, M. C., Gledhill, M., & Achterberg, E. P. (2011). Dissolve iron (III) speciation in the high latitude North Atlantic Ocean. Deep Sea Research. *Part 1: Oceanographic Reseach Papers, 58*(11), 1049-1059.

Mohamed, K. N., Rosli, N. S., Bohari, S. A., & Ismail, M. R. (2016). Fractionation of Mn, Fe and Cu in seawater at Port Dickson, Negeri Sembilan. In C. A. R. Mohamed (Ed.), In-house seminar of chemical oceanography (pp. 6-11). Bangi, Selangor: UKM.

Mokhtar, M. B., Aris, A. Z., Abdullah, M. H., Yusof, M. K., Abdullah, M. P., Idris, A. R., Raja Uzir, I. R. (2009). A pristine environment and water quality in perspective: Maliau Basin, Borneo's mysterious world. *Water and Environment Journal, 23*(3), 219-228.

Munksgaard, N. C., & Parry, D. L. (2001). Trace metals, arsenic and lead isotopes in dissolved and particulate phase of North Australian coastal and estuarine seawater. *Marine Chemistry, 75*(3), 165-184.

Neubauer, E., Köhler, S. J., vod der Kammer, F., Laudon, H., & Hofmann, T. (2013). Effect of pH and stream order on iron and arsenic speciation in boreal catchments. *Environmental Science and Technology, 47*(13), 7120-7128.

Nguyen, H. L., Leermakers, M., Elskens, M., Ridder, F. D., Doan, T. H., & Baeyens, W. (2005). Correlations, partitioning and bioaccumulation of heavy metals between different compartments of Lake Balaton. *Science of the Total Environment, 341*(2005), 211-226.

Pérez-Esteban, J., Escolástico, C., Moliner, A., & Masaguer, A. (2013). Chemical speciation and mobilization of copper and zinc in naturally contaminated mine soils with citric and tartaric acids. *Chemosphere, 90(2)*, 276-283.

Postma, H. (1967). Sediment transport and sedimentation in the estuarine environment. *American Association of Advanced Science, 83*(1967), 158-179.

Santana-Casiano, J. M., González-Dávila, M., & Millero, F. J. (2004). The oxidation of Fe (II) in NaCl–HCO3− and seawater solutions in the presence of phthalate and salicylate ions: a kinetic model. *Marine Chemistry, 85*(1-2), 27-40.

Schnoor, J. L. (1996). *Environmental modelling: Fate and transport of pollutants in water, air, and soil.* New York, USA: John Wiley & Sons.

Suratman, S., Hang H.C., & Shazili, N. A. M. (2009). A preliminary study of the distribution of selected trace metals in the Besut River basin, Terengganu, Malaysia. *Bulletin of Environmental Contamination and Toxicology, 82*(1), 16-19.

Syvitski, J. P. M., Kettner, A. J., D. Peckham, S. D., & Kao, S. J. (2005). Predicting the flux of sediment to the coastal zone: Application to the Lanyang Watershed, Northern Taiwan. *Journal of Coastal Research, 21*(3), 580-587.

United States Environmental Protection Agency. (2006). *Volunteer stream monitoring: A methods manual.* Office of Water, Washington D.C.: Author.

van der Berg, C. M., & Huang, Z. Q. (1984). Determination of iron in seawater using cathodic stripping voltammetry preceded by adsorptive collection with the hanging mercury drop electrode. *Journal of Electroanalytical Chemistry and Interfacial Electrochemistry, 177*(1-2), 269-280.

Velasquez, I., Nunn, B. L., Ibisanmi, E., Goodlet, D. R., Hunter, K. A., & Sander, S. G., (2011). Detection of hydroxamate siderophores in coastal and sub-antarctic waters off the southeastern coast of New Zealand. *Marine Chemistry 126*(1-4), 97-107.

Viers, J., Dupre, B., & Gaillardet, J. (2009). Chemical composition of suspended sediments in World Rivers: New insights from new database. *Science of the Total Environment, 407*(2), 853-868.

Watmough, S. A., Eimers, M. C., & Dillon, P. J. (2007). Manganese cycling in central Ontario forests: Response to soil acidification. *Applied Geochemistry, 22*(6), 1241-1247.

Xing, W., & G. Liu, G. (2011). Iron biogeochemistry and its environmental impacts in freshwater lakes. *Fresenius Environmental Bulletin, 20*(6), 1339-1345.

Yang, J. Y., Yang, X. E., He, Z. L., Li, T. Q., Shentu, J. L., & Stoffella, P. J. (2006). Effects of pH, organic acids, and inorganic ions on lead desorption from soils. *Environmental Pollution, 143*(1), 9-15.

Yap, C, K., Chee, M. W., Shamarina, S., Edward, F. B., Chew, W., & Tan, S. G. (2011). Penilaian kualiti air permukaan di perairan pantai Malaysia menggunakan analisis multivariat [Assessment of surface water quality in the Malaysian coastal waters by using multivariate analyses]. *Sains Malaysiana, 40*(10), 1053-1064.

Yusof, N., Haraguchi, A., Hassan, M. A., Othman, M. R., Wakisaka, M., & Shirai, Y. (2009). Measuring organic carbon, nutrients and heavy metals in rivers receiving leachate from controlled and uncontrolled municipal solid waste (MSW) landfills. *Waste Management, 29*(10), 2666-2680.

Zhang, D., Lee, D.-J. & Pan, X. (2013). Desorption of Hg (II) and Sb (V) on extracellular polymeric substances: Effects of pH, EDTA, Ca (II) and temperature shocks. *Bioresource Technology, 128*(2013), 711-715.

Zhang, D., Yuanxin, H., Xu, Z., Yuan, S., Li, Y., & Wang, L. (2017). Effect of overlying water pH, dissolved oxygen and temperature on heavy metal release from river sediments under laboratory conditions. *Archives of Environmental Protection, 43*(2), 28-36.

Pertanika J. Sci. & Technol. 27 (4): 2333 - 2349 (2019)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Hydrological Response of Semi-Arid River Catchment to Rainfall and Temperature Fluctuations

Da'u Abba Umar[1], Mohammad Firuz Ramli[2]*, Ahmad Zaharin Aris[2,3] and Muhammad Amar Zaudi[2]

[1]*Department of Environmental Sciences, Faculty of Science, Federal University Dutse, Dutse, Nigeria*
[2]*Faculty of Environmental Studies, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia*
[3]*Environmental Forensic Research Unit, Faculty of Environmental Studies, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia*

ABSTRACT

Determining the response of basin water resources to rainfall and temperature fluctuations is a crucial source of information for basins water resources planning and management. The study used a descriptive, Mann-Kendall trend test (M-K) and Multiple Linear Regression (MLR). The mean, standard deviations and variations were spatially interpolated using the geostatistical technique. The trend results showed an increase in both rainfall and temperature series. However, the only statistically significant trends were in June and September for rainfall series and in February, May, and April for the temperature series. Rainfall exhibited high temporal variability whereas temperature showed high spatial variability. The intra-annual variability was higher than the inter-annual variability, suggesting that the local climate is largely controlled by natural force. The result of the multiple linear regression (R^2 =0.431), indicates that the hydrology and water resources of the basin are impacted largely by factors not considered in this study such as land use changes, infiltration, and rate of evaporation among others. However, among the factor considered, rainfall (Beta = 0.505; P = 001) has the highest impacts on the river discharge behavior and should be given preference while addressing water resources predicaments in the catchment.

Keywords: Climate variability, geostatistical methods, multiple linear regression, river catchment, semi-arid region

ARTICLE INFO

Article history:
Received: 3 December 2018
Accepted: 30 January 2019
Published: 21 October 2019

E-mail addresses:
abbaumar.d@fud.edu.ng (Da'u Abba Umar)
firuz@upm.edu.my (Mohammad Firuz Ramli)
zaharin@upm.edu.my (Ahmad Zaharin Aris)
amarzaudi@gmail.com.my (Muhammad Amar Zaudi)
*Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

INTRODUCTION

Global advances in standard of living and increased economic development have resulted in an increasing demand and dependency on water resources. Many countries are projected to face water stress in the foreseeable future (Daniel et al., 2011). This forecasted water scarcity is associated with climate change and variability, increased human and animal population, urbanization and industrial development (Falkenmark, 2013). However, in semi-arid regions of the world, climate change and variability are identified as the most influential factor affecting water resources availability (Daniel et al., 2011). In these regions, our ability to clearly comprehend the current and future water resources availability depends largely on our understanding of the connections between climate variability in the form of rainfall and temperature fluctuations and the hydrological response of the basin ecosystem. This will help to appreciate surface water resources availability in response to climate variability.

Besides, rainfall and temperature changes, extreme weather events such as floods and droughts play an important role in the understanding of climatic changes. The historical beginning of the decrease in river discharge in Hadejia river catchment is dated back to 1970s and 1980s in response to the historical droughts that swept Sub-Saharan Africa (Umar et al., 2017). The consequence was the drastic effect incurred on the wet season flood/inundation used for flood recession agriculture in the area. The basin has experienced a decreasing trend in flood extent at the downstream areas from about 962 km^2 in 1991 to 525km^2 and 413 km^2 in 1992 and 1993 respectively (Adams & Thomas, 1996). Thus, from the foregoing, the need to assess the hydrological response of Hadejia river catchment to rainfall and temperature changes has arisen in a way to provide crucial information for effective planning and management of the basin's water resources.

MATERIALS AND METHODS

Study Area

Hadejia River Basin is a sub-catchment of the Hadejia-Jama'are river basin in the Northern Nigeria (Figure 1).

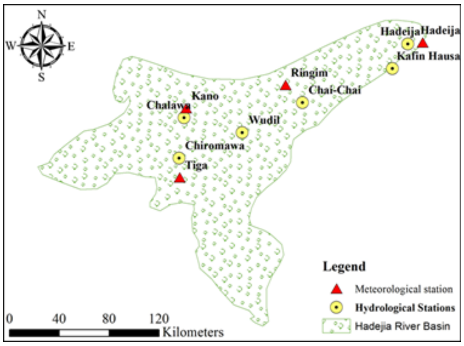


Figure 1. Hydrological and meteorological stations in the basin (Modified after IUCN 2003)

The Hadejia River was formed by the confluence of the Challawa and Kano Rivers at Tamburawa and it passed through two distinct geological formations, Basement Complex and Chad Formation. The elevation of the area reduces northward from over 600 m above sea level in the south to less than 300 m in the northern part of the basin (Adakayi, 2012).

Two wind systems control the climate of the area, the South West (SW) and North East (NE) trade winds. The SW trade winds come along with moisture from the Gulf of Gunue and stay in the North between May to September (summer), while the NE trade wind comes along with dry cold winds from the Sahara Desert and stays between Octobers to April (winter).

The mean annual rainfall varies from 1,100mm in the most southern locations of the basin (Tiga and Kano) to less than 300 mm in the extreme northeast (Hadejia). Temperatures reach as high as 35°C before the arrival of the rains (April/May) and drop as low as 18°C in December/January (Akinsanola & Ogunjobi, 2014).

Data

Rainfall, temperature and river discharge data for thirty-six years were obtained from Nigerian Meteorological Agency (NIMET), Hadejia Jama'are River Basin Development Authority (HJRBDA) and Jigawa State Ministry of Water Resources (JSMWR) respectively. These data were used for the analysis of hydroclimatic relationships and responses and only stations within the basin perimeters were used.

Prior to the application of suitable statistics, the data was subjected to QA/QC analysis (Duhan & Pandey, 2013), checked for irregularities such as missing data and outliers. The QA/QC scrutiny revealed that the data was statistically clean and normality distributed except for few missing data constituting less than 10% of the whole dataset and were replaced with the means of the last two recorded observations that binds the missing observations (Mwangi et al., 2016). For the normality test, the Shapiro-Wilk test was used to establish the normality status of data. The least square regression (LSR) test was, however, applied on only the rainfall and river discharge time series.

Data Analysis

Mann-Kendall (M-K) was used after the preliminary descriptive analysis. After assessing the temporal behavior of the data sets, multiple regression was conducted using rainfall, minimum and maximum temperature as independent variables (IV) and river discharge as the dependent variable (DV). The mean, SD and CV was interpolated using Inverse Distance Weighted (IDW), a geostatistical tool used for spatial analysis in the GIS environment

RESULTS AND DISCUSSIONS

Descriptive Statistics (DS) and Exploratory Data Analysis (EDA) of Rainfall Series

Prior to the statistical application normality test was conducted using Shapiro-Wilk normality test, the normality results for all the three sets of data (rainfall, temperature, and river discharge) was $p = 0.490$; $p = 0.319$; $p = 0.080$ respectively. Indicating that all the data set were normally distributed as represented in the Q-Q and Box plots in Figure 2a and 2f.

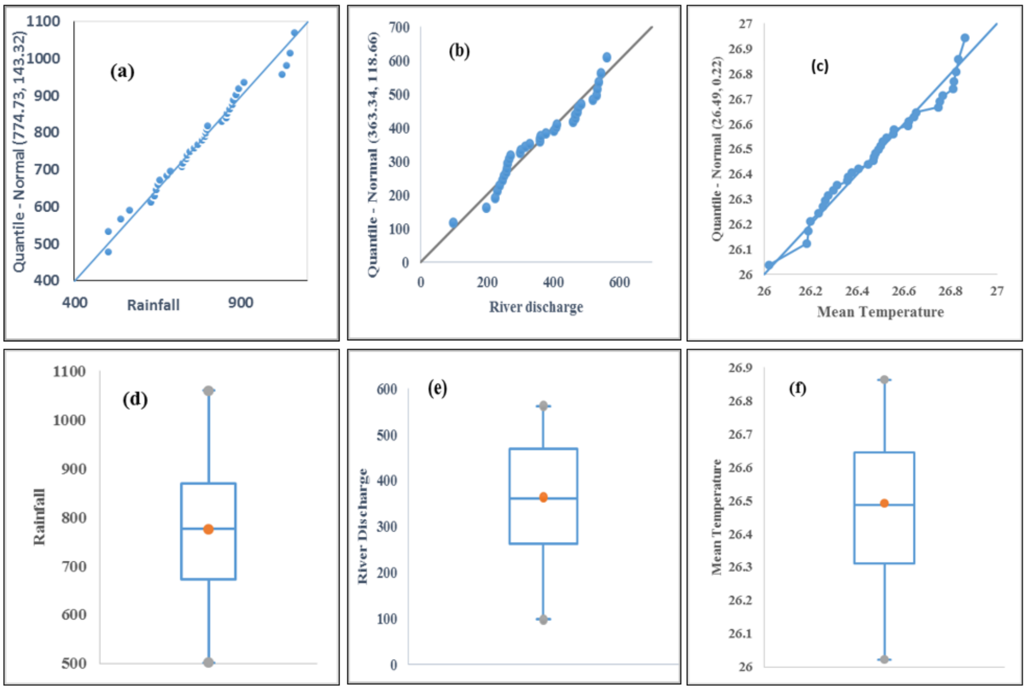


Figure 2. Normality distribution of annual mean rainfall (a) river discharge (b) temperature (c) and Q-Q plot of rainfall (d) river discharge (e) and temperature (f)

Additionally, the distribution of the rainfall, temperature and river discharge mean, SD and CV were spatially interpolated using IDW spatial interpolation tool (Figure 3(a) and 3(l)). Rainfall and river discharge shows a relatively similar pattern. The mean (Figure 3(a) and 3(d)) and standard deviation (Figure 3(e) and 3(h)) were found to be higher in the southern part of the basin and lower in the northern part. Meanwhile, the river discharge was slightly higher in the middle of the basin where the two rivers (Tiga and Challawa) converged to form the Hadejia river system at Wudil. However, the coefficient of variation exhibited a different pattern, where the higher CV was in the south-eastern part of the basin and the lowest CV was in the south-western part (Figure 3(i) and 3(l)).

In the case of temperature (Tmin and Tmax), the mean was higher in the northern part of the basin and decreased southward (Figure 3(b) and 3(c)). Meanwhile, SD and CV for Tmax were higher in the southern part of the basin (Figure 3(g) and 3(k)), while that of Tmin exhibited contrary pattern where SD and CV were higher in two different positions, extreme north and south of the basin (Figure 3(f) and 3(j)).

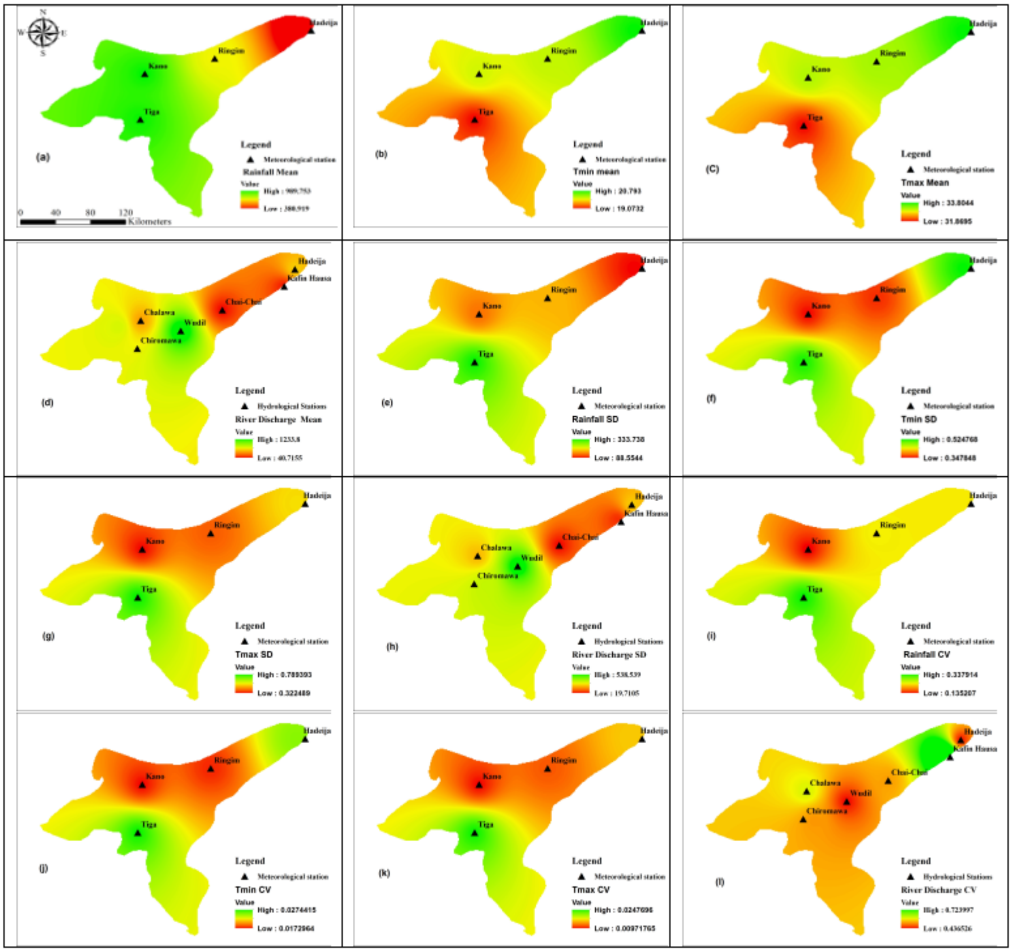


Figure 3. Spatial interpolation of mean (a) rainfall, (b) T_{min} (c) T_{max} and (d) river discharge; (e) rainfall SD (f) T_{min} SD (g) T_{max} and (h) river discharge; (i) rainfall CV (j) Tmin (k)T_{max} and (l) river discharge CV

The time series plots of rainfall, temperature, and river discharge showed that the mean annual rainfall and temperature were increasing over the entire basin (Table 1, 2, 3 and Figure 4). Meanwhile, mean annual river discharge was decreasing.

Table 1

Descriptive statistics and trend analysis (M-K) of rainfall for the entire basin

Stations	Minimum (mm/yr)	Maximum (mm/yr)	Mean (mm/yr)	SD (mm/yr)	CV (%)	Z (mm/yr)	Q (mm/yr)
Tiga	477.6	1789.2	987.6	333.7	33.0	1.51	0.746
Kano	700.8	1266.0	989.8	133.8	13.0	0.67	1.637
Ringim	398.4	1137.6	740.6	174.4	23.0	1.48	4.436
Hadejia	225.6	544.8	380.9	88.6	23.0	1.65	0.175
Basin	450.6	1184.4	774.7	182.6	20.0	1.32	1.74

Table 2

Descriptive statistics of monthly, seasonal and annual temperature

Series	T_{mean} Mean (°C)	T_{mean} SD (°C)	T_{mean} CV (%)	T_{max} Mean (°C)	T_{max} SD (°C)	T_{max} CV (%)	T_{min} Mean (°C)	T_{min} SD (°C)	T_{min} CV (%)
Jan	21.9	0.9	3.9	29.7	1.1	3.6	14.2	0.9	6.4
Feb	24.7	1.1	1.1	32.6	1.3	3.9	16.8	1.3	7.5
Mar	28.6	0.9	3.2	36.2	1.1	3.9	21.0	1.0	4.8
Apr	29.5	0.3	1.1	38.2	0.6	1.6	20.9	0.3	1.3
May	30.7	0.7	2.1	37.2	1.0	2.5	24.1	0.6	2.3
Jun	28.8	0.5	1.6	34.5	0.6	1.6	23.2	0.5	2.2
Jul	27.4	0.5	1.7	32.6	0.8	2.4	22.1	0.5	2.4
Aug	26.5	0.5	1.9	31.4	0.9	2.7	21.6	0.5	2.1
Sept	26.9	0.4	1.6	32.1	0.6	1.9	21.8	0.5	2.3
Oct	26.5	0.6	2.1	32.2	0.7	2.2	20.8	0.7	3.1
Nov	24.3	0.5	1.9	31.5	0.7	2.1	17.2	0.5	2.9
Dec	22.0	0.5	2.3	29.6	0.7	2.2	14.5	0.6	4.2
DJF	22.9	0.5	2.2	30.6	0.6	2.1	15.2	0.6	3.8
MAM	29.6	0.4	1.3	37.2	0.6	1.6	22.0	0.4	1.6
JJA	27.6	0.4	1.3	32.8	0.5	1.8	22.4	0.2	1.0
SON	25.9	0.4	1.4	31.9	0.4	1.4	19.8	0.4	1.9
Annual	26.5	0.2	0.8	33.1	0.3	0.9	19.8	0.2	1.0

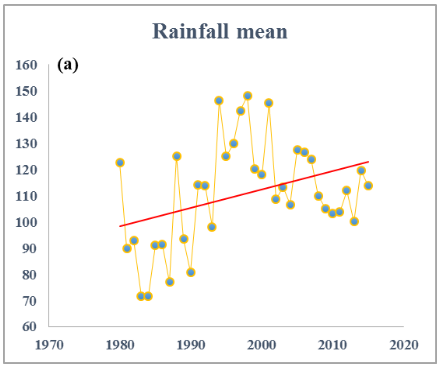
Note: SD and CV are deviation and standard coefficient of variability, respectively

Table 3

Descriptive statistics and trend analysis (M-K) of river discharge for the entire basin

Stations	Minimum (mm/yr)	Maximum (mm/yr)	Mean (mm/yr)	SD (mm/yr)	CV (%)	Z (mm/yr)	Q (mm/yr)
Chiromawa	205.0	788.4	473.1	187.9	40.0	0.50	1.558
Challawa	200.6	324.4	248.4	34.8	14.0	0.86	0.450
Wudil	348.8	1790.0	1085.6	440.8	41.0	0.11	0.236
Hadejia	172.5	877.0	412.4	195.9	47.0	-2.33	-7.087
Basin	218.2	682.4	483.6	188.8	25.0	-1.02	-1.356

Note: Confidence level at 95%



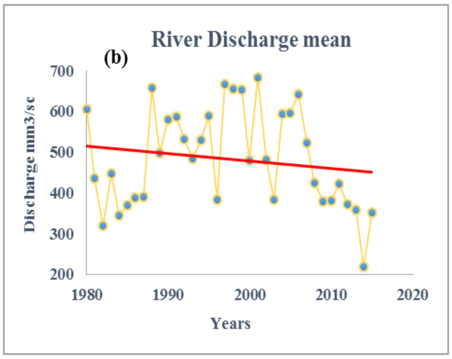


Figure 4. Time series plot of mean (a) rainfall (b) river discharge

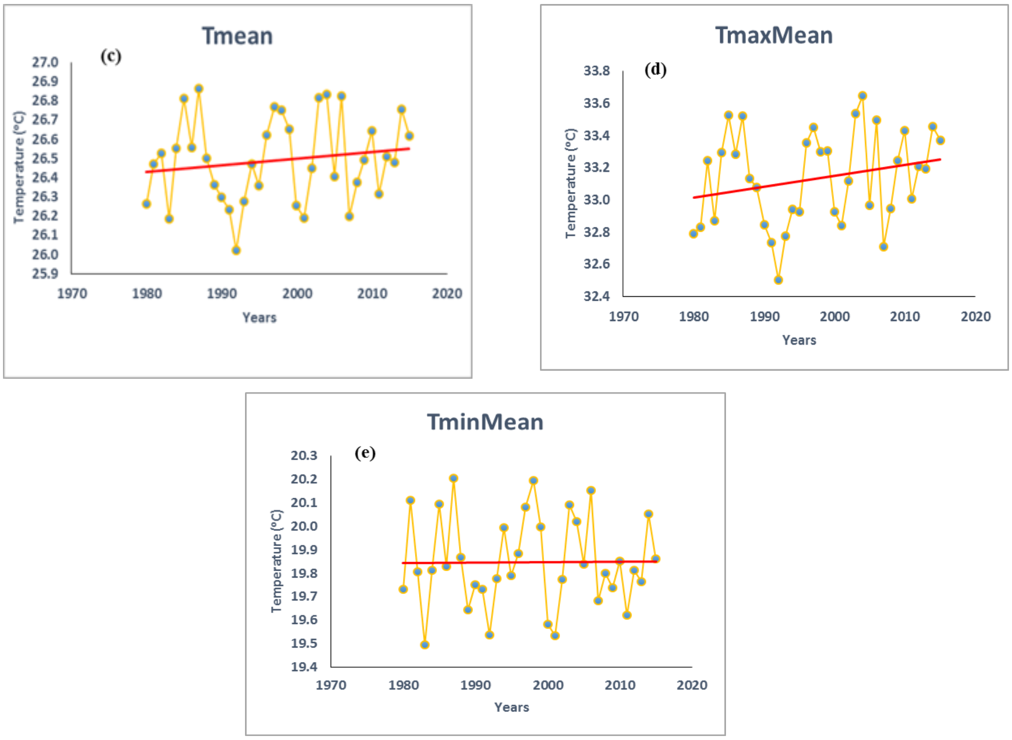


Figure 4. Time series plot of mean (c) temperature mean (d) maximum and (e) minimum temperature

The monthly rainfall, temperature, and river discharge trend analysis shows varied results of increasing and decreasing trends. For instance, noticeable increase in rainfall trend was detected except for the months of April and August, however, the statistically significant increasing trends were in June and September (Table 4).

Table 4

Descriptive statistics of monthly (mm/month) and annual rainfall series (mm/year)

Series	Minimum	Maximum	Mean	SD	CV (%)	Z	Q
April	0.0	38.9	17.3	11.1	63.4	-0.12	-0.042
May	21.0	109.7	57.4	24.0	41.2	0.75	0.234
June	57.0	158.9	114.4	27.4	23.6	**1.98**	0.933
July	102.5	333.1	207.4	49.3	23.5	1.73	1.367
Aug.	74.7	372.9	237.1	65.6	27.3	-0.20	-0.308
Sept.	38.5	243.7	124.7	43.6	34.5	**2.66**	1.563
Oct.	0.0	39.8	16.1	10.4	63.4	1.57	0.208
Annual	71.6	148.0	110.6	20.1	17.9	1.62	0.618

Note: Bold font is significant at 95%

However, the monthly temperature trends shows that T_{mean} was outweighed by negative trends and among the positively increasing months only February was statistically significant. Similarly, T_{min} displayed more negative trends, but only statistically significant in April, November, and December. However, the statistically significant positive trends in the T_{min} series were in July and August. Meanwhile, the T_{max} series was dominated by positive trends, but only February and May were statistically significant. The only statistically significant negative trends in T_{max} series was in August (Table 5).

Table 5

Trend statistics of monthly, seasonal and annual temperature series

Temperature Series	T_{mean}			T_{max}			T_{min}		
	Test Z	Sig.	Ch.	Test Z	Sig.	Ch.	Test Z	Sig.	Ch.
Jan	-0.25		-0.004	0.50		0.010	-1.20		-0.019
Feb	1.66	+	0.033	2.67	**	0.058	0.59		0.008
Mar	0.60		0.008	0.52		0.007	-0.03		0.000
Apr	-1.48		-0.008	0.98		0.010	-8.57	***	-0.026
May	0.71		0.008	1.81	+	0.038	-0.80		-0.009
Jun	-0.90		-0.007	-0.60		-0.007	-0.90		-0.009
Jul	1.27		0.013	-0.71		-0.012	3.12	**	0.025
Aug	-0.25		-0.001	-2.45	*	-0.030	2.89	**	0.023
Sept	0.08		0.000	-0.46		-0.003	1.66	+	0.013
Oct	-0.34		-0.004	-1.31		-0.016	1.16		0.012
Nov	-0.83		-0.007	0.54		0.008	-2.32	*	-0.020
Dec	-1.19		-0.010	0.78		0.013	-2.66	**	-0.025
DJF	0.79		0.006	2.38	*	0.026	-0.86		-0.009
MAM	1.10		0.009	2.04	*	0.022	-1.21		-0.008
JJA	-0.48		-0.003	-1.66	+	-0.017	6.52	***	0.015
SON	0.15		0.001	-0.19		-0.003	0.61		0.003
Annual	0.97		0.004	1.50		0.006	0.23		0.001

Notes: *** significant at $\alpha = 0.001$, ** significant at $\alpha = 0.01$, * significant at $\alpha = 0.05$, + significant at $\alpha = 0.1$ and - not significant. Ch. and Sig. refer to change and sign, respectively

On the other hand, the monthly river discharge trend results showed a general decreasing trend except for the months of June, November, and October. The statistically significant decreasing trend was in August (Table 6).

However, the spatial variations of river discharge trends across the basin showed varied results, where the upstream (Chiromawa & Challawa) and midstream (Wudil) stations displayed an insignificant increasing trend, while the downstream station (Hadejia) exhibited statistically significant decreasing trend (Table 3).

Table 6
Descriptive statistics of monthly (mm/month) and annual river discharge series (mm/year)

Series	Minimum	Maximum	Mean	SD	CV (%)	Z	Q
Jan.	3.3	15.0	11.0	2.6	24	-0.98	-0.045
Feb.	5.4	13.8	9.1	2.3	25	-1.10	-0.044
March	4.7	12.9	8.3	2.0	24	-0.29	-0.009
April	5.8	24.6	15.3	4.5	29	0.11	0.006
May	11.4	30.8	22.9	5.3	23	-1.59	-0.154
June	14.8	62.6	38.4	10.6	28	0.72	0.087
July	22.0	98.7	66.1	16.8	25	-1.51	-0.419
Aug.	29.5	151.0	97.0	28.1	29	**-1.73**	-0.741
Sept.	47.6	185.2	120.0	36.7	31	-0.59	-0.412
Oct.	14.6	139.2	59.4	33.2	56	0.00	0.009
Nov.	6.4	39.5	21.8	9.8	45	0.23	0.054
Dec.	5.5	22.4	14.1	4.3	31	-0.91	-0.085
Annual	218.2	682.4	483.4	118.8	25	-1.02	-1.356

Note: Bold font is significant at 95%

The summary of the descriptive statistic shows that annual minimum and maximum rainfalls were 450.6 and 1154.4 mm/yr, respectively. Mean rainfall was highest in August (237.1 mm/yr) and lowest in October (16.1 mm/yr) while the annual mean was 774.7 mm/yr. SD was highest in August (65.6 mm/yr) and lowest in October (10.4 mm/yr), with the annual SD obtained, was 20.1 mm/yr; CV was highest in April and October (63.4% each) and lowest in July (23.5%) whereas the annual CV was 17.9% (Table 4). The results show that rainfall variability was higher on intra-annual basis than inter-annual in the basin.

However, temperature descriptive summaries indicate that the spatial mean fluctuates from 21.9 to 30.7°C, SD from 0.3 to 1.1°C, maximum from 29.6°C to 38.2°C and minimum from 14.2 to 24.1°C (Table 5). The resultant warming due to temperature increase over the whole basin was relatively uniform, with CV ranging from 1.0% to 0.8%, but was 0.9% on the average. Higher values of temperature were obtained in the months of March and April (Table 5). However, the seasonal analysis showed that MAM (spring) had been the hottest of all seasons, while SON was the coldest, and monthly scales variability was higher compared to both seasonal and annual scale variability. The temporal plots of the cumulative departure of Tmean, Tmax, and Tmin (Figure 5) also showed similarity in their pattern of cooling and warming. For instance, between 1985 and 2005 there was a random warming in 1995. However, both Tmean and Tmin changed from warming to cooling in 2005, and continued till 2015 though there was an indication of reversal to the warming trend from that year.

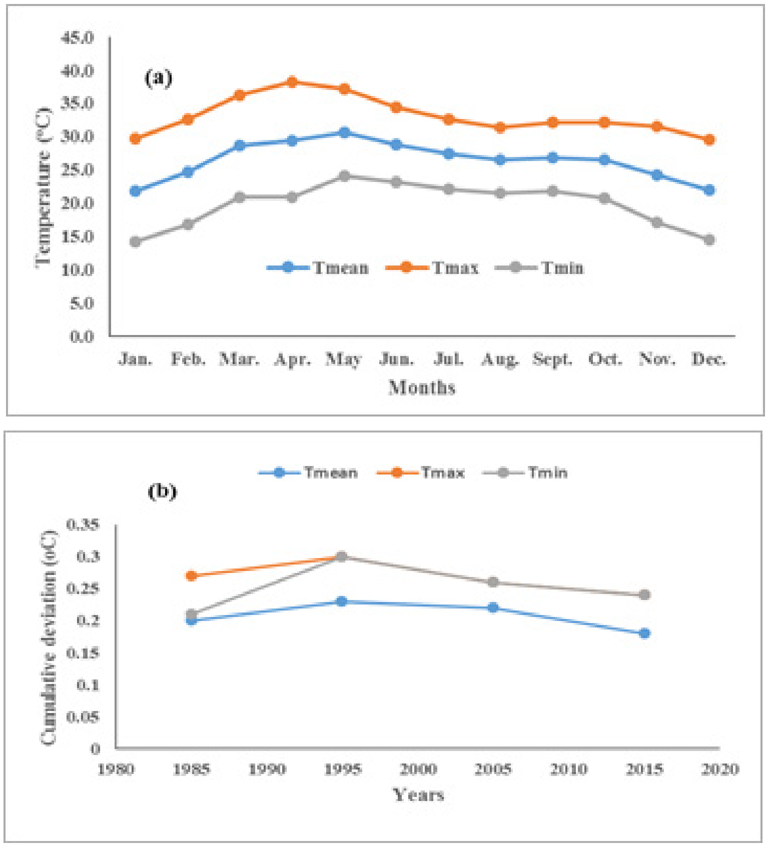


Figure 5. Plots of (a) monthly temperature series and (b) cumulative deviations

While the descriptive summaries of river discharge showed that annual minimum and maximum discharges were 218.2 and 682.4 mm^3/yr, respectively. Mean discharge was highest in September (120.0 mm^3/yr) and lowest in March (8.3 mm/yr) while the annual mean was 483.4 mm3/yr. SD was highest in September (36.7 mm^3/yr) and lowest in March (2.0 mm^3/yr), with the annual SD obtained as 118.8 mm3/yr, however, CV was highest in October (56%) and lowest in May (23%) and the annual CV was 25% (Table 4). The result shows that river discharge variability was higher on inter-annual than intra-annual basis within the basin.

Mann-Kendall (M-K)

The trend analysis test was conducted using the M-K test. The results for all the studied variables and the data points in the entire basin was communicated concurrently with the descriptive statistic results in Table 4, 5 and 6 for rainfall, temperature, and river discharge respectively. Rainfall annual trend showed a statistically insignificant increasing trend

(Table 4). The trend slope ranged from -0.042 to 0.933 mm/yr and for the entire basin it was 0.618 mm/yr. For the monthly analysis, it was also dominantly increasing trends except for April and August which showed insignificant decreasing trends. In all the positive trends only two monthly series (June and September) indicates a statistically significant trend with the slope of 0.933 mm/yr and 1563 mm/yr respectively (Table 4.0). Similarly, there were two other months (April and August) with insignificant decreasing trends, while the rest exhibited insignificant positive trends of varying degrees. The magnitude of change in the series ranged between -0.12 mm/yr (April) and 2.66 mm/yr (September). Although trends were dominantly positive in all the monthly and annual analysis except but in a few months, however, it seemed there was redistribution in the rainfall pattern, particularly in the monthly series.

Temperature trends result was summarized in Table 5. The result showed an annual increasing trend in both maximum and minimum temperature. The trend slope ranged from -0.042 to 0.933 mm/yr and for the entire basin it was 0.618 mm/yr. For the monthly analysis, it was also dominantly increasing trends except for April and August which showed insignificant decreasing trends.

Correlation and Regression analysis

Preliminary analysis via multiple correlation technique had shown that rainfall played positive roles in river discharge as could be inferred from their positive correlation coefficients (Rf; r = 0.539, n = 216, P = .001), while temperature (T_{min}; r = -0.034, n = 216, P = 0.353 and T_{max}; r = -0.315*, n= 216, P = 0.084) (Table 7) played negative role of reducing the river discharge.

Table 7

Correlations between RD (as DV) and Rf, T_{max} and T_{min} (as IDV)

Variables	RD	P-value
Rf	0.539**	0.001
Tmax	-0.315*	0.084
Tmin	-0.034	0.353

*Notes: ** Correlation is significant at the 0.01 level (2-tailed); * correlation is significant at the 0.05 level (2-tailed); RD=River discharge, T_{min} and T_{max} are the minimum and maximum temperature respectively.*

Accordingly, rainfall, maximum and minimum temperature was chosen for modeling the relationships between climate and river discharge. The relationship between the dependent variable (in this case river discharge) and the independent variables (rainfall and temperature) was examined through the following equation:

$$RD = a + b_1x_1 + b_2x_2 \ldots + b_nx_n + e \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots.. (6)$$

RRD= river discharge;
a = the intercept on the y- axis;
b_1 - b_n = partial regression coefficient of independent variables
x_1-x_n = the independent variable
e = random error of unexplained variables

The model was designed following Ekpoh and Ekpenyong (2011) in their study of the effects of climate variations on Water Yield in the Sokoto Region of Northern Nigeria. The climatic parameters chosen are expected to give details of their impact on river discharge in terms of their regression coefficients, while the P-values serve as the check on the 'goodness-of-fit' or otherwise of the relationships

The multiple regression model with all three predictors produced $R^2 = 0.431$ (Table 8). Rainfall had a statistically significant positive impact on the outcome of the dependent variable (river discharge). Thus, it was found to have significant positive regression weights (Table 8). However, where the maximum temperature contributed negatively to the multiple regression model, the minimum temperature contributed positively to the regression response variable, though statistically insignificant. Generally, the regression output has unveiled the significant role played by rainfall in the determination of river discharge characteristics in the catchments.

Table 8

Result of multiple regression analysis between Rainfall, Maximum and Minimum temperature and river

	a	Rainfall	Tmax	Tmin	R2	Adjusted R2
R Coefficient	0.656	0.505	-0.350	0.182	0.431	0.357
P-value	0.131	0.001	0.084	0.353		

Notes: Predictors: Rainfall, minimum and maximum temperature; Response variable: River discharge

Rainfall Analysis

Prior to the assessment of the individual impact of rainfall and temperature on the river discharge, trend analysis was conducted on both data series covering the period 1980 to 2015 over the HRB. The results showed that rainfall and temperature had increased over the entire basin, while river discharge was decreasing particularly downstream of the basin. Temporal variability was more noticeable than the spatial variability, and it was assumed the low spatial variability displayed might be due to the smaller spatial scope of the current study relative to some previous studies (e.g. Adakayi, 2012; Ifabiyi & Ojoye, 2013). On the other hand, annual rainfall variability does not differ much from the reported studies (Adewole & Serifat, 2015; Mohammed et al., 2015).

Meanwhile, the monthly analysis showed that rainfall was increasing in all the rain-bearing months except in April and August (Table 4), and only June and September were statistically significant. Furthermore, October was the driest month while August the wettest. Thus, peak discharge and flood events witnessed in the study area occured in the months of September largely due to August rainfall contribution and this scenario could continue for sometimes in the future. This result is in line with (Oyewole et al., 2014) who associated the devastating flood events in Nigeria with the increased in August and September rainfall.

Similarly, a noticeable increasing trend in rainfall was reported in the region by Adakayi (2012). He applied the M-K trend test to hydro-meteorological variables using 36 years of data in northern Nigeria where the HRB was located. In the same vein, Mertz et al. (2012) reported rainfall recovery in the Sub-Saharan West African beginning from the 1990s. Elsewhere in China, Qin et al. (2010) studied the impacts of climate changes on water resources at Tarim river basin, their results indicated an increasing trend in precipitation as shown by the current study. Similarly, Modarres and Silva (2007) reported an insignificant trend in annual rainfall in the semi-arid region of Iran.

Temperature trends results also showed a rising trend, and the rate of warming is relatively higher in T_{max} than in Tmin but the pattern was basically comparable. However, it appeared that the monthly temperature was highly variable as the CV was as high as 7.5% for T_{min}; while in the Tmax series it reached 3.9°C, SD went up to 1.3°C. This suggests that the monthly temperature behavior in the basin is more temporally inclined, with the highest rate in April and May but the trend is stronger in February and weaker in January and August, this is one of the characteristics of tropical semi-arid climate (Nobre et al., 2016). Consistent with the current findings, Adakayi (2012) reported a higher temperature in April and May. The April-May higher temperature recorded may perhaps be a warming up towards the summer (JJA) storms.

Furthermore, the magnitude of increased warming was found to be higher in Tmax than in both Tmean and Tmin indicating greater warming during the daytime relative to night-time. On the contrary, Oguntunde et al. (2012) discovered a night-time (Tmin) warming stronger than the day-time. A similar stronger significant trend in minimum temperature was reported over Italy (Brunetti et al., 2006), India (Kothawale et al., 2010) and Brazil (Nobre et al., 2016). The daytime increased warming experienced in the area, could be linked to the overriding influences of natural inherent climate variability.

However, the warming trends observed in the seasonal scales was only significant in winter and spring, for T_{mean} and T_{max} and in summer for the T_{min}, whereas there was a significant decreasing trend in summer for the T_{max} series. Moreover, the highest increased warming magnitudes were observed during JJA (summer) in the T_{min} series. Similar findings were reported in other parts of the world (e.g. Kothawale et al., 2010; Nobre et al., 2016). However, Gocic and Trajkovic (2013) reported contrary results where high-temperature trends were observed in autumn and spring seasons in Serbia. Though part of his findings can be sustained herewith since the current study has also observed significant warming trends in the spring season for the T_{mean} and T_{max} series. In the whole, the result could be ascribed to the temperature gradient that is typically created due to season's adjustment from warming to cooling, and that is yet not understood and further investigation is needed to unveil the possible reason(s) in future studies.

However, annual trends in all the temperature series were positively insignificant and magnitudes of change ranged between 0.001°C and 0.006°C concluding that the warming in the basin was not significant within the time span of the study. However, the result is consistent with some previous studies (Adakayi, 2012; Ayinde et al., 2013; Dammo et al., 2015).

Contrary to what was obtainable in both rainfall and temperature results, river discharge shows a decreasing annual trend attributed to many factors such as the climatic variations (rainfall and temperature fluctuations), the nature and the spatial disparity of the geology settings, high temperature and rate of evaporation downstream relative to other part of the basin. Moreover, there exist two non-returning channel (Chai Chai and Kafin Hausa bifurcation streams) before the outlet of the Hadejia river basin which occasionally reduces the volume of water reaching the downstream station.

Additionally, there is the disappearance of sharp channel streams at a certain location along the basin profile due to high erosion and siltation. Moreover, the basin management in itself was blamed for the reduced flow downstream of the basin, though some previous studies had disputed these allegations, stressing that, rainfall variability was the main factor determining the general flow of the river (Adewole & Serifat, 2015). Thus, of all the identified factors associated to decrease streamflow downstream, climate variability was conceived as the most frightening one.

The results of the regression analysis showed that the three predictors variables; rainfall, maximum and minimum temperature are not the only factors affecting the variations of the river discharge in Hadejia river basin, the evidence is the coefficient of determination obtained $R_2 = 0.432$, and the adjusted $R_2 = 0.357$ (Table 8). However, a close scrutiny of the Table reveals that rainfall made the highest contribution to the regression equation with a Beta coefficient of 0.505, which is highly significant at the 0.001 level. This is not surprising as some previous studies indicated rainfall as the most important parameters affecting river discharge variations, particularly in the semi-arid region of the country (Umar et al., 2017). On the other hand, both maximum and minimum temperature shows an inverse relationship with river discharge, however, with the insignificant negative Beta coefficient of -0.350 for Tmax and a positive insignificant Beta coefficient of 0.182 for the Tmin. The inference that could be made from the “cat and mouse” behavior of rainfall and temperature in relation to river discharge variations is that while rainfall increases river discharge also increases, but river discharge decreases with increasing temperature as the river water is subtracted via evaporation.

CONCLUSIONS

The findings showed that there was a strong positive relationship between rainfall and river discharge, suggesting that the river discharge was at the mercy of rainfall behavior. However, besides the predictor variables used in this study, there are other explanations associated with river discharge fluctuations in the basin such as the losses to ground water, evaporation and the water loses to the non-returning river channels (Chai-Chai and Kafin Hausa channels). Additionally, the two major dams constructed; Tiga and Challawa Gorge in 1974 and 1992 also affected the natural pattern of the Hadejia river flow turning the flow regime to perennial from the usual seasonal flows. In consideration of the decreasing trend of the river discharge at the downstream areas of the basin possible strategies are suggested such as the use of point-drop, pitcher or water-can irrigations in place of the sprinkler irrigation, the use of the soil surfaces mulching to reduce the water losses via evapotranspiration. However, water loses from the river surfaces can be reduced via aquifer recharge and recovery (ARR). Finally, massive afforestation exercise should be embarked upon in the HRB and indeed in the entire semi-arid region of Nigeria.

ACKNOWLEDGEMENTS

This paper is part of the project funded by Universiti Putra Malaysia (UPM) under the grant GP-IPS/2017/9574700. The first author want to thank the management of the Federal University Dutse (FUD), Dutse Nigeria and the Tertiary Education Trust Fund (TETFUND) Nigeria for sponsoring his PhD. program at the leading research university in Malaysia (UPM).

REFERENCES

Adakayi, P. E. (2012). *An assessment of rainfall and temperature variations in selected stations in parts of Northern Nigeria* (Doctoral thesis), University of Jos, Nigeria.

Adams, W. M., & Thomas, D. H. (1996).Conservation and sustainable resource use in the Hadejia–Jama’are Valley, Nigeria. *Oryx, 30*(2), 131-142.

Adewole, O. O., & Serifat, F. (2015). Modelling Rainfall Series in the Geo-Political Zones of Nigeria. *Journal of Environment Earth Science, 5*(2), 100-120.

Akinsanola, A., & Ogunjobi, K. (2014). Analysis of rainfall and temperature variability over Nigeria. *Global Journal of Human-Social Science Research, 14(*3), 1-19.

Ayinde, O., Ojehomon, V., Daramola, F., & Falaki, A. (2013). Evaluation of the effects of climate change on rice production in Niger state, Nigeria. *Ethiopian Journal of Environmental Studies and Management, 6*(6),763-73.

Brunetti, M., Maugeri, M., Monti, F., & Nanni, T. (2006). Temperature and precipitation variability in Italy in the last two centuries from homogenised instrumental time series. *International Journal of Climatology, 26*(3), 345-381.

Dammo, M., Ibn Abubakar, B., & Sangodoyin, A. (2015). Quantitative analysis of rainfall variations in north-eastern region of Nigeria. *Journal of Agricultural and Crop Research, 3*(5), 67-72.

Daniel, E. B., Camp, J. V., LeBoeuf, E. J., Penrod, J. R., Dobbins, J. P., & Abkowitz, M. D. (2011). Watershed modeling and its applications: A state-of-the-art review. *The Open Hydrology Journal, 5*(1), 26-50.

Duhan, D., & Pandey, A. (2013). Statistical analysis of long term spatial and temporal trends of precipitation during 1901-2002 at Madhya Pradesh, India. *Atmospheric Research, 122,* 136-149.

Ekpoh, I., & Ekpenyong, N. (2011). The effects of recent climatic variations on water yield in the Sokoto Region of Northern Nigeria. *International Journal of Business and Social Science, 2*(7), 251-256.

Falkenmark, M. (2013). Growing water scarcity in agriculture: Future challenge to global water security. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 371*(2002), 1-14.

Gocic, M., & Trajkovic, S. (2013). Analysis of changes in meteorological variables using Mann-Kendall and Sen's slope estimator statistical tests in Serbia. *Global and Planetary Change, 100,* 172-182.

Ifabiyi, I., & Ojoye, S. (2013). Rainfall trends in the Sudano-Sahelian ecological zone of Nigeria. *Earth Science Research, 2*(2), 194-202.

Kothawale, D., Revadekar, J., & Kumar, K. R. (2010). Recent trends in pre-monsoon daily temperature extremes over India. *Journal of Earth System Science, 119*(1), 51-65.

Mertz, O., D'haen, S., Maiga, A., Moussa, I. B., Barbier, B., Diouf, A., … Dabi, D. (2012). Climate variability and environmental stress in the Sudan-Sahel zone of West Africa. *Ambio, 41(*4), 380-392.

Modarres, R., & da Silva, V. d. P. R. (2007). Rainfall trends in arid and semi-arid regions of Iran. *Journal of Arid Environments, 70*(2), 344-355.

Mohammed, M. U., Abdulhamid, A., Badamasi, M., & Ahmed, M., (2015) Rainfall Dynamics and Climate Change in Kano, Nigeria. *Journal of Scientific Research & Reports 7*(5), 386-395.

Mwangi, H. M., Julich, S., Patil, S. D., McDonald, M. A., & Feger, K.-H. (2016). Relative contribution of land use change and climate variability on discharge of upper Mara River, Kenya. *Journal of Hydrology: Regional Studies, 5,* 244-260.

Nobre, C. A., Marengo, J. A., Seluchi, M. E., Cuartas, L. A., and Alves, L. M. (2016). Some characteristics and impacts of the drought and water crisis in Southeastern Brazil during 2014 and 2015. *Journal of Water Resource and Protection, 8*(02), 252-262.

Oguntunde, P. G., Abiodun, B. J., & Lischeid, G. (2012). Spatial and temporal temperature trends in Nigeria, 1901–2000. *Meteorology and Atmospheric Physics, 118*(1-2), 95-105.

Oyewole, J., Thompson, A., Akinpelu, J., & Jegede, O. (2014). Variation of rainfall and humidity in Nigeria. *Methodology, 4*(2014), 1-10.

Qin, B., Zhu, G., Gao, G., Zhang, Y., Li, W., Paerl, H. W., & Carmichael, W. W. (2010). A drinking water crisis in Lake Taihu, China: Linkage to climatic variability and lake management. *Environmental Management, 45*(1), 105-112.

Umar, D. A., Umar, H.A., & Tukur, A. I. (2017) Climate variability and water supply: A review of rural water planning techniques for semi-arid region of Nigeria. *Dutse Journal of Pure and Applied Sciences (DUJOPAS), 3*(2), 318-330.

Pertanika J. Sci. & Technol. 27 (4): 2351 - 2372 (2019)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Some Relation Properties of Rough Neutrosophic Multisets

Suriana Alias[1]*, Daud Mohamad[2] and Adibah Shuib[2]

[1]*Faculty of Computer and Mathematical Sciences, Universiti Teknologi Mara, 18500 UiTM, Machang, Kelantan, Malaysia*
[2]*Faculty of Computer and Mathematical Sciences, Universiti Teknologi Mara, 40450 UiTM, Shah Alam, Selangor, Malaysia*

ABSTRACT

Rough neutrosophic multisets are an improved model of generalization pf neutrosophic multisets represented within the Pawlak's boundary set of information: lower and upper approximation. The concepts of rough neutrosophic multisets can be easily extended to a relation, mainly since a relation is also a set, i.e. a subset of a Cartesian product. This paper establishes an axiomatic definition of rough neutrosophic multisets relation of Cartesian product over a universal set. Some of the operations and properties of rough neutrosophic multisets, such as max, min, the composition of two rough neutrosophic multisets relation and inverse rough neutrosophic multisets relation, are studied with a proven condition. An algorithm of rough neutrosophic multisets relation is also presented as a step followed to obtain the rough neutrosophic multisets relation. Successful analysis using rough neutrosophic multisets relation theory is represented by the illustrative example of expert opinion about automobile popularity. In conclusion, with a specified condition in uncertainty information, rough neutrosophic multiset relations are generalized in terms of the relation properties of a rough fuzzy relation, rough intuitionistic fuzzy relation, and rough neutrosophic relation over universal. Subsequently, their properties could also be examined.

Keywords: Max, min, neutrosophic multisets, relation, rough neutrosophic multisets, rough set, uncertainty

ARTICLE INFO

Article history:
Received: 29 Mar 2018
Accepted: 13 May 2019
Published: 21 October 2019

E-mail addresses:
suria588@uitm.edu.my (Suriana Alias)
daud@tmsk.uitm.edu.my (Daud Mohamad)
adibah@tmsk.uitm.edu.my (Adibah Shuib)
*Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

INTRODUCTION

The notion of relation plays a significant role in describing the relationship of strength and weakness between the element theory or variable presented in a universal set. For example, in the mathematical formulation of uncertainty information such as win and loss, and accept and reject conditions, the relation between these variables is defined to check whether there is a significant relationship among them. A bigger value for the relation indicates a stronger relationship between the information. This relation can be defined in a universal set or more than one universal set.

Uncertainty information always involves inconsistent information due to the results from opinion experts. The data set for uncertainty information is not suitable to be expressed by a standard mathematical data set. For example, an expert may give a different value for a person's intelligence, such as quite intelligent or more intelligent. These different values assigned to person intelligent is more valuable if there is a definition added by how many percentages the intelligent belong to that person. Inspired from that, fuzzy sets (FS) as defined by (Zadeh, 1965) formulated uncertainty information for inconsistent information. Zadeh introduced membership degree, assigned as (T_A (x)) to replace the inconsistent information with the membership degree value [0, 1]. By FS theory, the membership degree for quite intelligent was assigned as 0.4, and membership degree for more intelligent was assigned as 0.6. For this situation, especially for uncertain information, different experts might assign different values of a membership degree according to their judgement.

Later, Atanassov (1986) generalized FS to intuitionistic fuzzy set (IFS) where the uncertainty information involving incomplete and inconsistent information was overcome. He introduced a pair of value set memberships $(T_A(x), F_A(x))$, namely membership degree ($T_A(x)$) and non-membership degree ($F_A(x)$) in an uncertain information environment. The values of the membership degree and non-membership degree are also in the interval [0, 1]. FS and IFS theory are modelled based on the successful mathematical formulation for incomplete and inconsistent information but are still lacking in formulating indeterminate information. Occasionally, when dealing with uncertain information, there is still a need for indeterminate information. For example, in a voting system, the percentage for who have voted and who has not voted can be represented by IFS, but in voting, we must also consider the result for those who have not voted at all (indeterminate).

In IFS theory, this indeterminate information is represented by $1\text{-}T_A$ (x)-F_A (x). But it is still not a concrete value for indeterminate value, since the value can be zero. Therefore, Smarandache (1998) initiated the theory of neutrosophic set (NS) and neutrosophic logic (NL) by generalizing FS, IFS, fuzzy logic (FL) and intuitionistic fuzzy logic (IFL). Compared to FS and IFS, NS can express the indeterminate information by introduced the membership degree of indeterminacy (I_A (x)) and adding in incomplete and inconsistent information. The triple set member (T_A (x),I_A (x),F_A (x)) is defined for NS. The value for membership degree, indeterminacy degree and non-membership degree also lies between [0,1].

In general, more than one or repeated uncertainty information also take into consideration in formulating the uncertainty model. For example, in medical analysis, the different time is taken (morning, noon, evening) of the symptoms such as temperature and headache of the patient will help to get an accurate diagnosis. Later, for the first time, Smarandache (2013) refined NS to a neutrosophic refined set (NRS). The element set of NS in NRS is allowed for repeated with different or same elements; (T1, T2, …, T_m and I1, I2, …, In and F1, F2, …, Fr). Then, single valued neutrosophic multisets (SVNM) (Ye & Ye, 2014) inspired by NRS (Smarandache, 2013) was published with operation and properties of NS number. Instead of one-time for each element, the NRS and SVNM also allowed an element to occur more than once with possibly the same or different truth membership sequences, indeterminacy membership sequences, and falsity membership sequences $({T_A}^i(x), {F_A}^i(x), {I_A}^i(x))$. The combination idea of the NRS and SVNM is used in our research.

Recently, Alias et al. (2017) introduced the concept of rough neutrosophic multiset (RNM) which was a hybrid of the rough set (RS) and SVNM theory for solving multiple uncertain information in Pawlak's approximation. The RNM deals with truth membership sequence, indeterminate membership sequence, and falsity membership sequence. The thoery is a generalization of the RNS and from interest of researchers to formulate the uncertain and imprecise information. It has become a valuable component in investigating decision-making analysis, pattern and selection recognition, and the control issue. For example, in selection system for university acceptance (accept, neutral, reject) of multiple choice of candidate, there must be multiple phases of the selection process in university before the best candidate is chosen. All the processes in the selection system will be formulated by RNM theory.

This paper introduces the RNM relation by first giving their basic operations and properties of relation set based on RNM theory. The RNM relation is studied to verify the interaction between variables used in the RNM theory, such as types of car with the relation of popularities of car in making the right decision to buy or sold it. RNM relation is another new tool in dealing with uncertain and imprecise information.

This paper is structured by first presenting the axiomatic definition of RNM relation in the universe set as a novel notion. Before that, important preliminaries related to RNM relation is presented. Next, some relation of RNM relation such as max, min, the composition between two RNM relation, and the inverse of RNM relation are studied. Subsequently, their properties is also examined. The illustrative example with algorithm is presented to verify the RNM relation application. Then, the results and discussions section is given. Lastly, the conclusion section provides an overall summary of RNM relation performance.

MATERIAL AND METHOD

Generalization of Rough Neutrosophic Multisets Relation

This section recalls the important definition to define the rough neutrosophic multisets (RNM) relation. All the proofs of operations and properties may be obtained from (Alias et al., 2017; Nguyen et al., 2014; Pawlak, 1982; Ye & Ye, 2014). Some basic concepts of uncertainty information and hybrid uncertainty information are also discussed in this section.

Some Basic Concepts of Uncertainty Information and Hybrid Uncertainty Information

Many theories were later introduced with the same direction of a fuzzy sets, such as soft set (SS) (Molodtsov, 1999) and rough set (RS) (Pawlak, 1982). The hybrid mathematical formulation of the neutrosophic set (NS) and neutrosophic multisets (NM) for solving uncertainty information has been widely explored. Mandal (2015) combined the theory of NS with SS and introduced the Neutrosophic Soft Set (NSS) to solve the uncertainty involving soft set condition. Then, Alkhazaleh (2016) generalized NSS by introducing the concept of time-neutrosophic soft set and study some of its properties. Ali & Smarandache, (2017) introduced the Complex Neutrosophic Set (CNS) by hybridization of complex set and NS, and successfully introduced the relation and operation of CNS. Broumi et al. (2014) combined the RS and NS by introducing the Rough Neutrosophic Set (RNS). The theory of RNS is well discussed and applied in decision making (Mondal & Pramanik, 2015; Mondal et al., 2016; Pramanik & Mondal, 2015; Yang et al., 2017). Then, Broumi and Smarandache (2014) introduced interval valued Neutrosophic Rough Set (IVNRS) to overcome the interval set of RNS. In any case, every part of these theories has its specialities and difficulties.

Definition 1 (Rough set). Let R be an equivalence relation on the universal set V. Then, the pair (V,R) is called a Pawlak's approximation space. An equivalence class of R containing element set $\varkappa$ will be denoted by $[\varkappa]_R$. Now, for $X \subseteq V$, the upper and lower approximation of X with the respect to (*U,R*) are denoted by, respectively A1 ($\varkappa$) and A2 ($\varkappa$), and are defined as:

$$A_1(\varkappa) = \{\varkappa : [\varkappa]_{\mathcal{R}} \subseteq X\} \text{ and } A_2(\varkappa) = \{\varkappa : [\varkappa]_{\mathcal{R}} \cap X = \emptyset\} \qquad (1)$$

Now, if $A_1(\varkappa)=A_2(\varkappa)$ then X is called definable; otherwise, the pair $A(X) = (A_1(\varkappa), A_2(\varkappa))$ is called the rough set of X in U.

Some concepts of Single Valued Neutrosophic Multisets (SVNM) and Rough Neutrosophic Multisets (RNM)

Single valued neutrosophic multisets (SVNM) (Ye & Ye, 2014) inspired from NRS (Smarandache, 2013), and generalization of Fuzzy multiset (FM) (Yager, 1986), and Intuitionistic Fuzzy Multisets (Shinoj & John, 2012) was published with operation and properties of neutrosophic number.

Definition 2 (Single valued Neutrosophic Multisets). Let E be a universe. A neutrosophic multisets NM(H) on E can be defined in Equation 2 as:

$$NM(\mathcal{H}) = \{<\varkappa, \left(T_{\mathcal{H}}^{1}(\varkappa), T_{\mathcal{H}}^{2}(\varkappa), \ldots, T_{\mathcal{H}}^{p}(\varkappa)\right), \left(I_{\mathcal{H}}^{1}(\varkappa), I_{\mathcal{H}}^{2}(\varkappa), \ldots, I_{\mathcal{H}}^{p}(\varkappa)\right),$$

$$(F_{\mathcal{H}}^{1}(\varkappa), F_{\mathcal{H}}^{2}(\varkappa), \ldots, F_{\mathcal{H}}^{p}(\varkappa))>: \varkappa \in E\} \tag{2}$$

where, $T_{\mathcal{H}}^{1}(\varkappa), T_{\mathcal{H}}^{2}(\varkappa), \ldots, T_{\mathcal{H}}^{p}(\varkappa)$ is the truth membership sequence, $\left(I_{\mathcal{H}}^{1}(\varkappa), I_{\mathcal{H}}^{2}(\varkappa), \ldots, I_{\mathcal{H}}^{p}(\varkappa)\right.$ is the indeterminacy membership sequence, $\left(F_{\mathcal{H}}^{1}(\varkappa), F_{\mathcal{H}}^{2}(\varkappa), \ldots, F_{\mathcal{H}}^{p}(\varkappa)\right.$ is the falsity membership sequence, *p* is dimension of the *NM(H)*, $\varkappa$ is element set in universe *E*, and *H* is NM element in universe *E*.

The triple membership sequences of NM may be in a decreasing or increasing order. For convenience, the *NM(H)* can be denoted as:

$$NM(\mathcal{H}) = \{<\varkappa, \left(T_{\mathcal{H}}^{i}(\varkappa), I_{\mathcal{H}}^{i}(\varkappa), F_{\mathcal{H}}^{i}(\varkappa)\right)>: \varkappa \in E, i = 1, 2, \ldots, p\} \tag{3}$$

where i=1,2,…,p is NM sequence order, $T_{\mathcal{H}}^{i}(\varkappa)$ is the truth membership sequence, $I_{\mathcal{H}}^{i}(\varkappa)$ is the indeterminacy membership sequence, and $F_{\mathcal{H}}^{i}(\varkappa)$ is the falsity membership sequence. Also, $T_{\mathcal{H}}^{i}(\varkappa), I_{\mathcal{H}}^{i}(\varkappa), F_{\mathcal{H}}^{i}(\varkappa) \in [0,1]$ satisfies the condition

$$0 \leq T_{\mathcal{H}}^{i}(\varkappa) + I_{\mathcal{H}}^{i}(\varkappa) + F_{\mathcal{H}}^{i}(\varkappa) \leq 3 \text{ for any } \varkappa \in E \text{ and } i = 1, 2, \ldots, p.$$

Definition 3 (Rough Neutrosophic Multisets). Let E be a non-null set and R be an equivalence relation on E. Let H be neutrosophic multisets (NM) in E with the truth-membership sequence $(T_{\mathcal{H}}^{i})$, indeterminacy-membership sequences $(I_{\mathcal{H}}^{i})$ and falsity-membership sequences (F_H^i). The lower and the upper approximations of H in the approximation (E,R) denoted by $\underline{Nm}$(H) and $\overline{\boldsymbol{Nm}}(\mathcal{H})$ are respectively defined in Equation 4 as:

$$\underline{\boldsymbol{Nm}}(\mathcal{H}) = \left\{\langle \boldsymbol{x}, \left(\boldsymbol{T}_{\underline{\boldsymbol{Nm}}(\mathcal{H})}^{\boldsymbol{i}}(\boldsymbol{x}), \boldsymbol{I}_{\underline{\boldsymbol{Nm}}(\mathcal{H})}^{\boldsymbol{i}}(\boldsymbol{x}), \boldsymbol{F}_{\underline{\boldsymbol{Nm}}(\mathcal{H})}^{\boldsymbol{i}}(\boldsymbol{x}),\right)\rangle \middle| \boldsymbol{y} \in [\boldsymbol{x}]_{\boldsymbol{R}}, \boldsymbol{x} \in \boldsymbol{E}\right\}, \text{ and}$$

$$\overline{\boldsymbol{Nm}}(\mathcal{H}) = \left\{\langle \boldsymbol{x}, \left(\boldsymbol{T}_{\overline{\boldsymbol{Nm}}(\mathcal{H})}^{\boldsymbol{i}}(\boldsymbol{x}), \boldsymbol{I}_{\overline{\boldsymbol{Nm}}(\mathcal{H})}^{\boldsymbol{i}}(\boldsymbol{x}), \boldsymbol{F}_{\overline{\boldsymbol{Nm}}(\mathcal{H})}^{\boldsymbol{i}}(\boldsymbol{x}),\right)\rangle \middle| \boldsymbol{y} \in [\boldsymbol{x}]_{\boldsymbol{R}}, \boldsymbol{x} \in \boldsymbol{E}\right\} \tag{4}$$

where $\boldsymbol{T}_{\underline{\boldsymbol{Nm}}(\mathcal{H})}^{\boldsymbol{i}}(\boldsymbol{x}) = \wedge_{\boldsymbol{y} \in [\boldsymbol{x}]_{\boldsymbol{R}}} \boldsymbol{T}_{\mathcal{H}}^{\boldsymbol{i}}(\boldsymbol{y})$, $\boldsymbol{I}_{\underline{\boldsymbol{Nm}}(\mathcal{H})}^{\boldsymbol{i}}(\boldsymbol{x}) = \vee_{\boldsymbol{y} \in [\boldsymbol{x}]_{\boldsymbol{R}}} \boldsymbol{I}_{\mathcal{H}}^{\boldsymbol{i}}(\boldsymbol{y})$, $\boldsymbol{F}_{\underline{\boldsymbol{Nm}}(\mathcal{H})}^{\boldsymbol{i}}(\boldsymbol{x}) = \vee_{\boldsymbol{y} \in [\boldsymbol{x}]_{\boldsymbol{R}}} \boldsymbol{F}_{\mathcal{H}}^{\boldsymbol{i}}(\boldsymbol{y})$,

$\boldsymbol{T}_{\overline{\boldsymbol{Nm}}(\mathcal{H})}^{\boldsymbol{i}}(\boldsymbol{x}) = \vee_{\boldsymbol{y} \in [\boldsymbol{x}]_{\boldsymbol{R}}} \boldsymbol{T}_{\mathcal{H}}^{\boldsymbol{i}}(\boldsymbol{y})$, $\boldsymbol{I}_{\overline{\boldsymbol{Nm}}(\mathcal{H})}^{\boldsymbol{i}}(\boldsymbol{x}) = \wedge_{\boldsymbol{y} \in [\boldsymbol{x}]_{\boldsymbol{R}}} \boldsymbol{I}_{\mathcal{H}}^{\boldsymbol{i}}(\boldsymbol{y})$, $\boldsymbol{F}_{\overline{\boldsymbol{Nm}}(\mathcal{H})}^{\boldsymbol{i}}(\boldsymbol{x}) = \wedge_{\boldsymbol{y} \in [\boldsymbol{x}]_{\boldsymbol{R}}} \boldsymbol{F}_{\mathcal{H}}^{\boldsymbol{i}}(\boldsymbol{y})$.

Here i=1,2,…,p and positive integers are NM sequence order, ∧ and ∨ denote "min" and "max" operators, respectively; [x]_R is the equivalence class of the x, $T_{\underline{Nm}(\mathcal{H})}^{i}(x)$ is a lower approximation of NM truth membership sequence, $I_{\underline{Nm}(\mathcal{H})}^{i}(x)$ is the lower approximation of the NM indeterminacy membership sequence, $F_{\underline{Nm}(\mathcal{H})}^{i}(x)$ is the approximation of the NM falsity membership sequence, $\mathrm{T}_{\overline{\mathrm{Nm}}(\mathcal{H})}^{\mathrm{i}}(\mathrm{x})$ is the upper approximation of NM truth membership sequence, $I^{i}_{\overline{Nm}(\mathcal{H})}(x)$ is the approximation of NM indeterminacy membership sequence, $F^{i}_{\overline{Nm}(\mathcal{H})}(x)$ is the upper approximation of NM falsity membership sequence, $\boldsymbol{T^{i}_{\mathcal{H}}(y), I^{i}_{\mathcal{H}}(y)}$ **and** $\boldsymbol{F^{i}_{\mathcal{H}}(y)}$ and $F^{i}_{\overline{Nm}(\mathcal{H})}(x)$ are the membership sequences, indeterminacy sequences, and non-membership sequences of y with respect to H.

It can be said that $\boldsymbol{T^{i}_{\underline{Nm}(\mathcal{H})}(x), I^{i}_{\underline{Nm}(\mathcal{H})}(x), F^{i}_{\underline{Nm}(\mathcal{H})}(x)} \in [0,1]$ and $\boldsymbol{0 \leq T^{i}_{\underline{Nm}(\mathcal{H})}(x) + I^{i}_{\underline{Nm}(\mathcal{H})}(x) + F^{i}_{\underline{Nm}(\mathcal{H})}(x) \leq 3}$. Then, $\underline{Nm}(H)$ is a NM. Similarly, we have $\boldsymbol{T^{i}_{\overline{Nm}(\mathcal{H})}(x), I^{i}_{\overline{Nm}(\mathcal{H})}(x), F^{i}_{\overline{Nm}(\mathcal{H})}(x)} \in [0,1]$ and $0 \leq \boldsymbol{T^{i}_{\overline{Nm}(\mathcal{H})}(x) + I^{i}_{\overline{Nm}(\mathcal{H})}(x) + F^{i}_{\overline{Nm}(\mathcal{H})}(x)} \leq 3$. Then, $\overline{Nm}(\mathcal{H})$ is NM. Since $\underline{Nm}$ (H) and $\overline{Nm}(\mathcal{H})$ are two NMs in E, the NM mappings $\underline{Nm}, \overline{Nm}: Nm(E) \rightarrow Nm(E)$ are respectively referred to as the lower and upper NM approximation operators, while the pair of $\left(\underline{Nm}(\mathcal{H}), \overline{Nm}(\mathcal{H})\right)$ are called the rough neutrosophic multisets (RNM) in (U,R), respectively.

Rough Neutrosophic Multisets Relation

In this section, the Cartesian product of two rough neutrosophic multisets (RNM) over a universe is defined for RNM theory. Then, the RNM relation is examined for their desired properties. Lastly, the illustrative example followed the algorithm given is presented.

Assumptions and Notations

The RNM relation has been developed based on the following assumptions and notations.

Assumption. In the following section, we have considered only the case of the same repeated numbers of components 1,2,…,p for truth membership sequences $({T_{\mathcal{H}}}^{i})$, indeterminacy membership sequences $({I_{\mathcal{H}}}^{i})$, and falsity membership sequences $({F_{\mathcal{H}}}^{i})$ of NM H. We use NM numbers in the example of RNM relation.

Notations. The following notations as shown in Table 1 are used in the RNM relation model.

Table 1

Notations and descriptions used in the rough neutrosophic multiset relation model

Notation		Description
Q	:	Non-empty/ universe set
T	:	Truth membership degree
I	:	Indeterminacy membership degree
F	:	Falsity membership degree

Table 1 *(Continued)*

Notation		Description
X, Y, Z, Z'	:	Rough neutrosophic multiset in Q
h	:	Element set X in universe Q
g	:	Element set Y in universe Q
$\imath$	:	Element set Z in universe Q
t	:	Element set Z' in universe Q
$\varkappa, \hbar$	:	Element set in universe Q
i	:	Integer iteration number for neutrosophic multisets
p	:	Integer sequence for neutrosophic multisets
$Q \times Q$	:	Cartesian product of element set in universe Q
$X \times Y$	:	Cartesian product of X and Y in universe Q
$Y \times Z$	:	Cartesian product of Y and Z in universe Q
$X \times Z$	:	Cartesian product of X and Z in universe Q
$Z \times Z'$	:	Cartesian product of Z and Z' in universe Q
$Y \times X$	:	Cartesian product of Y and X in universe Q
(h, g)	:	A pair of elements set h and g in the universe relation $Q \times Q$
$(h, \imath)$	:	A pair of elements set h and $\imath$ in the universe relation $Q \times Q$
$(\imath, h)$	:	A pair of elements set $\imath$ and h in the universe relation $Q \times Q$
$(g, \imath)$	:	A pair of elements set g and $\imath$ in the universe relation $Q \times Q$
$(\imath, g)$	:	A pair of elements set $\imath$ and g in the universe relation $Q \times Q$
(h, t)	:	A pair of elements set h and t in the universe relation $Q \times Q$
(g, t)	:	A pair of elements set g and t in the universe relation $Q \times Q$
$(\imath, t)$	:	A pair of elements set $\imath$ and t in the universe relation $Q \times Q$
(g, h)	:	A pair of elements set g and h in the universe relation $Q \times Q$
$T_X^i(h)$	:	Truth membership sequence for X
$T_Y^i(g)$	:	Truth membership sequence for Y
$I_X^i(h)$	:	Indeterminacy membership sequence for X
$I_Y^i(g)$	:	Indeterminacy membership sequence for Y
$F_X^i(h)$	:	Falsity membership sequence for X
$F_Y^i(g)$	:	Falsity membership sequence for Y
$T_{X\times Y}^i(h, g)$	:	Truth membership sequence for cartesian product of X and Y
$I_{X\times Y}^i(h, g)$	:	Indeterminacy membership sequence for cartesian product of X and Y

Table 1 *(Continued)*

Notation		Description
$F^{i}_{X\times Y}(h, g)$	:	Falsity membership sequence for cartesian product of X and Y
$\mathfrak{R}$	:	Rough neutrosophic multisets relation on $Q \times Q$ based on the $X \times Y$
$\mathbf{T}^{i}_{\mathfrak{R}}(h, g)$	:	Truth-membership sequence relation for (h, g)
$\boldsymbol{I}^{i}_{\mathfrak{R}}(h, g)$	:	Indeterminacy-membership sequence relation for (h, g)
$\boldsymbol{F}^{i}_{\mathfrak{R}}(h, g)$	:	Falsity-membership sequence relation for (h, g)
$\mathcal{R}$	:	Equivalence relation of neutrosophic multisets
$\mathcal{R}_Q$	:	Equivalence relation of neutrosophic multisets in universe Q
$\underline{X \times Y}$	:	Lower approximation of neutrosophic multisets $X \times Y$ in universe Q
$\overline{X \times Y}$	:	Upper approximation of neutrosophic multisets $X \times Y$ in universe Q
$\underline{X \times Z}$	:	Lower approximation of neutrosophic multisets $X \times Z$ in universe Q
$\overline{X \times Z}$	:	Lower approximation of neutrosophic multisets $X \times Z$ in universe Q
$\underline{\mathcal{R}_Q}(X)$	:	Equivalence relation of lower approximation neutrosophic multisets X in universe Q
$\underline{\mathcal{R}_Q}(Y)$	:	Equivalence relation of lower approximation neutrosophic multisets Y in universe Q
$\underline{\mathcal{R}_Q}(X) \times \underline{\mathcal{R}_Q}(Y)$	:	Cartesian product of lower approximation $\underline{\mathcal{R}_Q}(X)$ and $\underline{\mathcal{R}_Q}(Y)$ in universe Q
$\overline{\mathcal{R}}_Q(X)$	:	Equivalence relation of upper approximation neutrosophic multisets X in universe Q
$\overline{\mathcal{R}}_Q(Y)$	:	Equivalence relation of upper approximation neutrosophic multisets Y in universe Q
$\overline{\mathcal{R}}_Q(X) \times \overline{\mathcal{R}}_Q(Y)$	:	Cartesian product of upper approximation $\underline{\mathcal{R}_Q}(X)$ and $\underline{\mathcal{R}_Q}(Y)$ in universe Q
a, b, c, d	:	Constant value for rough neutrosophic multisets relation in between [0, 1]
$M(\mathfrak{R})$	:	Matrix form for rough neutrosophic multisets relation
$M(\mathfrak{R})^{t}$	:	Transpose matrix form for rough neutrosophic multisets relation
$\mathfrak{R}_1, \mathfrak{R}_2, \mathfrak{R}_3$	:	Rough neutrosophic multisets relation on $Q \times Q$ based on the $X \times Y$
$\mathfrak{R}^{-1}$	:	Inverse rough neutrosophic multisets relation on $Q \times Q$ based on the $X \times Y$
$T^{i}_{\mathfrak{R}_1}(h, g)$, $T^{i}_{\mathfrak{R}_2}(h, g)$	:	Truth membership sequence for rough neutrosophic multisets relation of (h, g)

Table 1 *(Continued)*

Notation		Description
$I^i_{\Re_1 \wedge \Re_2}(h, g)$	:	Indeterminacy membership sequence for intersection operation of rough neutrosophic multisets relation of (h, g)
$F^i_{\Re_1 \wedge \Re_2}(h, g)$	:	Falsity membership sequence for intersection operation of rough neutrosophic multisets relation of (h, g)
$\Re_1 \vee \Re_2$	:	Union operator of two rough neutrosophic multisets relation
$T^i_{\Re_1 \vee \Re_2}(h, g)$	:	Truth membership sequence for union operation of rough neutrosophic multisets relation of (h, g)
$I^i_{\Re_1 \vee \Re_2}(h, g)$	:	Indeterminacy membership sequence for union operation of rough neutrosophic multisets relation of (h, g)
$F^i_{\Re_1 \vee \Re_2}(h, g)$	:	Falsity membership sequence for union operation of rough neutrosophic multisets relation of (h, g)
$\Re_1 \otimes \Re_2$	:	Multiply operator of two rough neutrosophic multisets relation
$T^i_{\Re_1 \otimes \Re_2}(h, g)$	:	Truth membership sequence for multiply operation of rough neutrosophic multisets relation of (h, g)
$I^i_{\Re_1 \otimes \Re_2}(h, g)$	:	Indeterminacy membership sequence for multiply operation of rough neutrosophic multisets relation of (h, g)
$F^i_{\Re_1 \otimes \Re_2}(h, g)$	:	Falsity membership sequence for multiply operation of rough neutrosophic multisets relation of (h, g)
$\Re_1 \oplus \Re_2$	:	Addition operator of two rough neutrosophic multisets relation
$T^i_{\Re_1 \oplus \Re_2}(h, g)$	:	Truth membership sequence for addition operation of rough neutrosophic multisets relation of (h, g)
$I^i_{\Re_1 \oplus \Re_2}(h, g)$	:	Indeterminacy membership sequence for addition operation of rough neutrosophic multisets relation of (h, g)
$F^i_{\Re_1 \oplus \Re_2}(h, g)$	:	Falsity membership sequence for addition operation of rough neutrosophic multisets relation of (h, g)
$\Re_1 \circ \Re_2$	:	Composition operator of two rough neutrosophic multisets relation
$\Re_1 \circ (\Re_2 \circ \Re_3)$	:	Composition operator of three rough neutrosophic multisets relation
$T^i_{\Re_1 \circ \Re_2}(h, \imath)$	:	Truth membership sequence for composition operation of rough neutrosophic multisets relation of $(h, \imath)$
$I^i_{\Re_1 \circ \Re_2}(h, \imath)$	:	Indeterminacy membership sequence for composition operation of rough neutrosophic multisets relation of $(h, \imath)$
$F^i_{\Re_1 \circ \Re_2}(h, \imath)$	:	Falsity membership sequence for composition operation of rough neutrosophic multisets relation of $(h, \imath)$

Table 1 *(Continued)*

Notation		Description
$T^{i}_{\Re_2}(g,\imath)$	:	Truth membership sequence for rough neutrosophic multisets relation of $(g,\imath)$
$I^{i}_{\Re_2}(g,\imath)$	:	Indeterminacy membership sequence for rough neutrosophic multisets relation of $(g,\imath)$
$F^{i}_{\Re_2}(g,\imath)$	:	Falsity membership sequence for rough neutrosophic multisets relation of $(g,\imath)$
$T^{i}_{\Re_1\circ(\Re_2\circ\Re_3)}(h,t)$	:	Truth membership sequence for composition operation of rough neutrosophic multisets relation of (h,t)
$T^{i}_{(\Re_2\circ\Re_3)}(g,t)$	:	Truth membership sequence for composition operation of rough neutrosophic multisets relation of (g,t)
$T^{i}_{\Re_3}(\imath,t)$	:	Truth membership sequence for composition operation of rough neutrosophic multisets relation of $(\imath,t)$
$T^{i}_{\Re^{-1}}(g,h)$	:	Truth membership sequence for inverse rough neutrosophic multisets relation of (g,h)
$I^{i}_{\Re^{-1}}(g,h)$	:	Indeterminacy membership sequence for inverse rough neutrosophic multisets relation of (g,h)
$F^{i}_{\Re^{-1}}(g,h)$	:	Falsity membership sequence for inverse rough neutrosophic multisets relation of (g,h)
$(\Re_1\circ\ \Re_2)^{-1}$	:	Inverse composition operator for rough neutrosophic multisets relation
${\Re_1}^{-1}, {\Re_2}^{-1}$	:	Inverse relation for rough neutrosophic multisets relation
$T^{i}_{(\Re_1\circ\Re_2)^{-1}}(\imath,h)$	:	Truth membership sequence for inverse composition operation of rough neutrosophic multisets relation of $(\imath,h)$
$T^{i}_{(\Re_2)^{-1}}(\imath,g)$	:	Truth membership sequence for inverse operation of rough neutrosophic multisets relation of $(\imath,g)$
$h\mathcal{R}_Q g$	:	Equivalence relation of (h,g) in universe Q

Phase 1: Development of Operation and Properties for Rough Neutrosophic Multisets Relation Theory

Definition 4. Let Q be a non-empty set and X and Y be the rough neutrosophic multiset (RNM) in Q. Then, the Cartesian product of X and Y is RNM in $Q\times Q$, denoted by $X\times Y$, is defined as:

$$X\times Y=\left\{<(h,g),\left(T^{i}_{X\times Y}(h,g)\right)\left(I^{i}_{X\times Y}(h,g)\right),\left(F^{i}_{X\times Y}(h,g)\right)>:(h,g)\in Q\times Q\right\} \quad (5)$$

where

$T^i_{X\times Y}(h,g) = \min\{T^i_X(h), T^i_Y(g)\}$, $I^i_{X\times Y}(h,g) = \max\{I^i_X(h), I^i_Y(g)\}$,

$F^i_{X\times Y}(h,g) = \max\{F^i_X(h), F^i_Y(g)\}$, $T^i_{X\times Y}, I^i_{X\times Y}, F^i_{X\times Y}: Q \to [0,1]$, and $i = 1, 2, \dots, p$.

Definition 5. Let Q be a non-empty set and X and Y be the rough neutrosophic multisets (RNM) in Q. We call $\mathfrak{R} \subseteq Q \times Q$ is RNM relation on Q×Q based on the X×Y, where X×Y is characterized by the truth-membership sequence $\mathbf{T}^{i}_{\mathfrak{R}}$, the indeterminacy-membership sequences $\mathbf{I}^{i}_{\mathfrak{R}}$ and the falsity-membership sequences $\mathrm{F}^{i}_{\mathfrak{R}}$, defined as:

$$\mathfrak{R} = \left\{< (h,g), \left(T^i_{\mathfrak{R}}(h,g)\right)\left(I^i_{\mathfrak{R}}(h,g)\right), \left(F^i_{\mathfrak{R}}(h,g)\right) >: (h,g) \in Q \times Q\right\} \qquad (6)$$

with conditions if satisfied:

(1) i) $T^i_{\mathfrak{R}}(h,g) = 1$ for all $(h,g) \in \underline{X \times Y}$ where $\underline{X \times Y} = \underline{\mathcal{R}_Q}(X) \times \underline{\mathcal{R}_Q}(Y)$,

ii) $T^i_{\mathfrak{R}}(h,g) = 0$, for all $(h,g) \in Q \times Q - \overline{X \times Y}$ where $\overline{X \times Y} = \overline{\mathcal{R}}_Q(X) \times \overline{\mathcal{R}}_Q(Y)$,

iii) $0 < T^i_{\mathfrak{R}}(h,g) < 1$, for all $(h,g) \in \overline{X \times Y} - \underline{X \times Y}$.

(2) i) $I^i_{\mathfrak{R}}(h,g) = 0$, for all $(h,g) \in \underline{X \times Y}$ where $\underline{X \times Y} = \underline{\mathcal{R}_Q}(X) \times \underline{\mathcal{R}_Q}(Y)$,

ii) $I^i_{\mathfrak{R}}(h,g) = 1$, for all $(h,g) \in Q \times Q - \overline{X \times Y}$ where $\overline{X \times Y} = \overline{\mathcal{R}}_Q(X) \times \overline{\mathcal{R}}_Q(Y)$,

iii) $0 < I^i_{\mathfrak{R}}(h,g) < 1$, for all $(h,g) \in \overline{X \times Y} - \underline{X \times Y}$.

(3) i) $F^i_{\mathfrak{R}}(h,g) = 0$, for all $(h,g) \in \underline{X \times Y}$ where $\underline{X \times Y} = \underline{\mathcal{R}_Q}(X) \times \underline{\mathcal{R}_Q}(Y)$,

ii) $F^i_{\mathfrak{R}}(h,g) = 1$, for all $(h,g) \in Q \times Q - \overline{X \times Y}$ where $\overline{X \times Y} = \overline{\mathcal{R}}_Q(X) \times \overline{\mathcal{R}}_Q(Y)$,

iii) $0 < F^i_{\mathfrak{R}}(h,g) < 1$, for all $(h,g) \in \overline{X \times Y} - \underline{X \times Y}$.

The RNM relation can be presented by relational tables and matrices like FS and IFS relation representation. Since the triple (T^i_A, I^i_A, F^i_A) has values within the interval [0,1], the elements of the neutrosophic matrix also have values within [0,1].

Now, we can consider some properties of RNM relation. All the properties must satisfy the RNM definition and RNM relation condition in Definitions 4 and 5, respectively.

Proposition 1. Let $\mathfrak{R}_1, \mathfrak{R}_2$ be two rough neutrosophic multisets relation on $Q\times Q$ based on the X×Y. Then $\mathfrak{R}_1 \wedge \mathfrak{R}_2$, where

$$T^i_{\mathfrak{R}_1\wedge\mathfrak{R}_2}(h,g) = \min\{T^i_{\mathfrak{R}_1}(h,g), T^i_{\mathfrak{R}_2}(h,g)\},$$
$$I^i_{\mathfrak{R}_1\wedge\mathfrak{R}_2}(h,g) = \max\{I^i_{\mathfrak{R}_1}(h,g), I^i_{\mathfrak{R}_2}(h,g)\},$$
$$F^i_{\mathfrak{R}_1\wedge\mathfrak{R}_2}(h,g) = \max\{F^i_{\mathfrak{R}_1}(h,g), F^i_{\mathfrak{R}_2}(h,g)\}$$

for all $(h, g) \in Q \times Q$, is a rough neutrosophic multisets on Q×Q based on the X×Y and i=1,2,…,p.

Proof We show that $\mathfrak{R}_1 \wedge \mathfrak{R}_2$ satisfies Definition 5.

1) i) Since $T^i_{\mathfrak{R}_1}(h, g) = T^i_{\mathfrak{R}_2}(h, g) = 1$ for all $(h, g) \in \underline{X \times Y}$ then

$$T^i_{\mathfrak{R}_1 \wedge \mathfrak{R}_2}(h, g) = \min\{T^i_{\mathfrak{R}_1}(h, g), T^i_{\mathfrak{R}_2}(h, g)\} = 1 \text{ for all } (h, g) \in \underline{X \times Y}.$$

ii) Since $T^i_{\mathfrak{R}_1}(h, g) = T^i_{\mathfrak{R}_2}(h, g) = 0$ for all $(h, g) \in Q \times Q - \overline{X \times Y}$

then $T^i_{\mathfrak{R}_1 \wedge \mathfrak{R}_2}(h, g) = \min\{T^i_{\mathfrak{R}_1}(h, g), T^i_{\mathfrak{R}_2}(h, g)\} = 0$ for all

$(h, g) \in Q \times Q - \overline{X \times Y}$.

iii) Since $0 < T^i_{\mathfrak{R}_1}(h, g), T^i_{\mathfrak{R}_2}(h, g) < 1$, for all $(h, g) \in \overline{X \times Y} - \underline{X \times Y}$ then

$0 < T^i_{\mathfrak{R}_1 \wedge \mathfrak{R}_2}(h, g) = \min\{T^i_{\mathfrak{R}_1}(h, g), T^i_{\mathfrak{R}_2}(h, g)\} < 1$ for all

$(h, g) \in \overline{X \times Y} - \underline{X \times Y}$.

A similar proofing step is satisfied for conditions 2 and 3 in definition 5.

Proposition 2. Let $\mathfrak{R}_1, \mathfrak{R}_2$ be two rough neutrosophic multisets relation on *Q×Q* based on the X×Y. Then $\mathfrak{R}_1 \vee \mathfrak{R}_2$, where

$$T^i_{\mathfrak{R}_1 \vee \mathfrak{R}_2}(h, g) = \max\{T^i_{\mathfrak{R}_1}(h, g), T^i_{\mathfrak{R}_2}(h, g)\},$$

$$I^i_{\mathfrak{R}_1 \vee \mathfrak{R}_2}(h, g) = \min\{I^i_{\mathfrak{R}_1}(h, g), I^i_{\mathfrak{R}_2}(h, g)\},$$

$$F^i_{\mathfrak{R}_1 \vee \mathfrak{R}_2}(h, g) = \min\{F^i_{\mathfrak{R}_1}(h, g), F^i_{\mathfrak{R}_2}(h, g)\}$$

for all $(h, g) \in Q \times Q$, is a rough neutrosophic multisets on Q×Q based on the *X×Y* and i=1,2,…,p.

Proof We show that $\mathfrak{R}_1 \vee \mathfrak{R}_2$, satisfies definition 5.

(1) i) Since $T^i_{\mathfrak{R}_1}(h, g) = T^i_{\mathfrak{R}_2}(h, g) = 1$ **for all** $(h, g) \in \underline{X \times Y}$ **then**

$$T^i_{\mathfrak{R}_1 \vee \mathfrak{R}_2}(h, g) = \max\{T^i_{\mathfrak{R}_1}(h, g), T^i_{\mathfrak{R}_2}(h, g)\} = 1 \textbf{ for all } (h, g) \in \underline{X \times Y}.$$

ii) Since $T^i_{\mathfrak{R}_1}(h, g) = T^i_{\mathfrak{R}_2}(h, g) = 0$ for all $(h, g) \in Q \times Q - \overline{X \times Y}$

then $T^i_{\mathfrak{R}_1 \vee \mathfrak{R}_2}(h, g) = \max\{T^i_{\mathfrak{R}_1}(h, g), T^i_{\mathfrak{R}_2}(h, g)\} = 0$ for all $(h, g) \in Q \times Q - \overline{X \times Y}$.

iii) Since $0 < T^i_{\mathfrak{R}_1}(h, g), T^i_{\mathfrak{R}_2}(h, g) < 1$, for all $(h, g) \in \overline{X \times Y} - \underline{X \times Y}$ then

$0 < T^i_{\mathfrak{R}_1 \vee \mathfrak{R}_2}(h, g) = \max\{T^i_{\mathfrak{R}_1}(h, g), T^i_{\mathfrak{R}_2}(h, g)\} < 1$ for all $(h, g) \in \overline{X \times Y} - \underline{X \times Y}$.

A similar proofing step is satisfied for conditions 2 and 3 in definition 5.

Lemma 1. If $0<\varkappa,\hbar<1$, then

$0<\varkappa\hbar<1$ (Obvious)

$0<\varkappa+\hbar-\varkappa\hbar<1$.

Since $0<\varkappa,\hbar<1$ then $\varkappa+\hbar\geq 2\sqrt{\varkappa\hbar}>2\varkappa\hbar>\varkappa\hbar>0$, therefore $\varkappa+\hbar-\varkappa\hbar>0$. On the other hand, $1-(\varkappa+\hbar-\varkappa\hbar)=(1-\varkappa)(1-\hbar)>0$ then $\varkappa+\hbar-\varkappa\hbar<1$.

The following properties of RNM relation are obtained using these algebraic results (Lemma 1)

Proposition 3. Let $\mathfrak{R}_1, \mathfrak{R}_2$ be two rough neutrosophic multisets relations on Q×Q based on the X×Y. Then $\mathfrak{R}_1 \otimes \mathfrak{R}_2$, where,

$$T^i_{\mathfrak{R}_1 \otimes \mathfrak{R}_2}(h, g) = T^i_{\mathfrak{R}_1}(h, g) \cdot T^i_{\mathfrak{R}_2}(h, g),$$

$$I^i_{\mathfrak{R}_1 \otimes \mathfrak{R}_2}(h, g) = I^i_{\mathfrak{R}_1}(h, g) + I^i_{\mathfrak{R}_2}(h, g) - I^i_{\mathfrak{R}_1}(h, g) \cdot I^i_{\mathfrak{R}_2}(h, g)$$

$$F^i_{\mathfrak{R}_1 \otimes \mathfrak{R}_2}(h, g) = F^i_{\mathfrak{R}_1}(h, g) + F^i_{\mathfrak{R}_2}(h, g) - F^i_{\mathfrak{R}_1}(h, g) \cdot F^i_{\mathfrak{R}_2}(h, g)$$

for all $(h, g) \in Q \times Q$, is a rough neutrosophic multisets on Q×Q based on the X×Y and i=1,2,…,p.

Proof The relation $\mathfrak{R}_1 \otimes \mathfrak{R}_2$, satisfies definition 5.

(1) i) Since $T^i_{\mathfrak{R}_1}(h, g) = T^i_{\mathfrak{R}_2}(h, g) = 1$ **for all** $(h, g) \in \underline{X \times Y}$ **then**

$T^i_{\mathfrak{R}_1 \otimes \mathfrak{R}_2}(h, g) = T^i_{\mathfrak{R}_1}(h, g) \cdot T^i_{\mathfrak{R}_2}(h, g) = 1$ **for all** $(h, g) \in \underline{X \times Y}$**.**

ii) Since $T^i_{\mathfrak{R}_1}(h, g) = T^i_{\mathfrak{R}_2}(h, g) = 0$ for all $(h, g) \in Q \times Q - \overline{X \times Y}$

then $T^i_{\mathfrak{R}_1 \otimes \mathfrak{R}_2}(h, g) = T^i_{\mathfrak{R}_1}(h, g) \cdot T^i_{\mathfrak{R}_2}(h, g) = 0$ for all

$(h, g) \in Q \times Q - \overline{X \times Y}$.

iii) Since $0 < T^i_{\mathfrak{R}_1}(h, g), T^i_{\mathfrak{R}_2}(h, g) < 1$, for all $(h, g) \in \overline{X \times Y} - \underline{X \times Y}$ then

$0 < T^i_{\mathfrak{R}_1 \otimes \mathfrak{R}_2}(h, g) = T^i_{\mathfrak{R}_1}(h, g) \cdot T^i_{\mathfrak{R}_2}(h, g) < 1$ for all $(h, g) \in \overline{X \times Y} - \underline{X \times Y}$

(Lemma 1 (i)).

The proof is also true for conditions 2 and 3 by following Lemma 1 (ii).

Proposition 4. Let $\mathfrak{R}_1, \mathfrak{R}_2$ be two rough neutrosophic multisets relations on *Q*×*Q* based on the *X*×*Y*. Then $\mathfrak{R}_1 \otimes \mathfrak{R}_2$, where,

$$T^i_{\mathfrak{R}_1 \oplus \mathfrak{R}_2}(h,g) = T^i_{\mathfrak{R}_1}(h,g) + T^i_{\mathfrak{R}_2}(h,g) - T^i_{\mathfrak{R}_1}(h,g) \cdot T^i_{\mathfrak{R}_2}(h,g),$$

$$I^i_{\mathfrak{R}_1 \oplus \mathfrak{R}_2}(h,g) = I^i_{\mathfrak{R}_1}(h,g) \cdot I^i_{\mathfrak{R}_2}(h,g),$$

$$F^i_{\mathfrak{R}_1 \oplus \mathfrak{R}_2}(h,g) = F^i_{\mathfrak{R}_1}(h,g) \cdot F^i_{\mathfrak{R}_2}(h,g)$$

for all $(h,g) \in Q \times Q$, *is a rough neutrosophic multisets on Q × Q based on the X × Y and i=1,2,…,p.*

Proof The relation satisfies $\mathfrak{R}_1 \otimes \mathfrak{R}_2$, definition 5 and Lemma 1. All the proving steps are similar to Proposition 3.

Phase 2: Development of the Composition of Two Rough Neutrosophic Multisets Relation Theory

The composition of two relations is essential for applications in real life. This relation can be computed over a universe with useful significance. For example, in the same group of medical diagnoses, the relation between patient and symptoms, and the relation between disease and symptoms, will be composited to find the relation between patient and disease. Some of the applications of the composition of two relation sets were discussed by Guo et al. (2017) and Yang et al. (2016).

Definition 6. Let *Q* be a non-empty set and X, Y and Z be the RNM in *Q*. Let $\mathfrak{R}_1, \mathfrak{R}_2$ are two RNM relations on *Q*×*Q*, based on *X*×*Y*,*Y*×*Z*, respectively. The composition of $\mathfrak{R}_1, \mathfrak{R}_2$ denoted as $\mathfrak{R}_1 \circ \mathfrak{R}_2$ as defined by Q×Q based on X×Z, where,

$$T^i_{\mathfrak{R}_1 \circ \mathfrak{R}_2}(h,\imath) = \max_{g \in Q}\{\min[T^i_{\mathfrak{R}_1}(h,g), T^i_{\mathfrak{R}_2}(g,\imath)]\},$$

$$I^i_{\mathfrak{R}_1 \circ \mathfrak{R}_2}(h,\imath) = \min_{g \in Q}\{\max[I^i_{\mathfrak{R}_1}(h,g), I^i_{\mathfrak{R}_2}(g,\imath)]\},$$

$$F^i_{\mathfrak{R}_1 \circ \mathfrak{R}_2}(h,\imath) = \min_{g \in Q}\{\max[F^i_{\mathfrak{R}_1}(h,g), F^i_{\mathfrak{R}_2}(g,\imath)]\}$$

for all $(h,\imath),(g,\imath) \in Q \times Q$ and i=1,2,…,p.
All the propositions must satisfy the condition in definition 4 and 5, respectively.

Proposition 5. $\mathfrak{R}_1 \circ \mathfrak{R}_2$ is a rough neutrosophic multisets relation on Q×Q based on X×Z.
Proof Since R_1,R_2 are two rough neutrosophic multisets relations on Q×Q based on X×Y,Y×Z, respectively:

Similar proving steps follow for conditions 2 and 3.

(1) **i)** **Then** $T^i_{\Re_1}(h,\imath) = 1 = T^i_{\Re_2}(h,\imath)$ **for all** $(h,\imath) \in \underline{X \times Z}$**. Let** $(h,\imath) \in \underline{X \times Z}$**, now**

$T^i_{\Re_1 \circ \Re_2}(h,\imath) = \max_{g \in Q}\{\min[T^i_{\Re_1}(h,g), T^i_{\Re_2}(g,\imath)]\} = 1$**. This holds for all** $(h,\imath) \in \underline{X \times Z}$**.**

ii) Let $(h,\imath) \in Q \times Q - \overline{X \times Z}$. So, $T^i_{\Re_1}(h,\imath) = 0 = T^i_{\Re_2}(h,\imath)$ for all $(h,\imath) \in Q \times Q - \overline{X \times Z}$. Then

$T^i_{\Re_1 \circ \Re_2}(h,\imath) = \max_{g \in Q}\{\min[T^i_{\Re_1}(h,g), T^i_{\Re_2}(g,\imath)]\} = 0$ for all $(h,\imath) \in Q \times Q - \overline{X \times Z}$.

iii) Again, since $0 < T^i_{\Re_1}(h,\imath), T^i_{\Re_2}(h,\imath) < 1$, for all $(h,\imath) \in \overline{X \times Z} - \underline{X \times Z}$,

then $0 < max_{g \in Q}\{\min[T^i_{\Re_1}(h,g), T^i_{\Re_2}(g,\imath)]\} < 1$ such that $0 < T^i_{\Re_1 \circ \Re_2}(h,\imath) < 1$ for all $(h,\imath) \in$

$\overline{X \times Z} - \underline{X \times Z}$.

Proposition 6. Let Q be a non-empty set. R_1,R_2,R_3 are rough neutrosophic multisets relations on Q×Q based on X×Y,Y×Z,Z×Z', respectively. Then $(\Re_1 \circ \Re_2) \circ \Re_3 = \Re_1 \circ (\Re_2 \circ \Re_3)$. The proof is similar to indeterminacy function and falsity function.

Proof For all $h, g, \imath, t \in Q$ we have

$$T^i_{\Re_1 \circ (\Re_2 \circ \Re_3)}(h,t) = \max_{g \in Q}\{\min[T^i_{\Re_1}(h,g), T^i_{(\Re_2 \circ \Re_3)}(g,t)]\}$$

$$= \max_{g \in Q}\{\min[T^i_{\Re_1}(h,g), \max_{\imath \in Q}\{\min[T^i_{\Re_2}(g,\imath), T^i_{\Re_3}(\imath,t)]]\}\}$$

$$= \max_{\imath \in Q}\{\min\{\max_{g \in Q}\{\min[T^i_{\Re_1}(h,g), T^i_{\Re_2}(g,\imath)], T^i_{\Re_3}(\imath,t)]\}\}$$

$$= \max_{\imath \in Q}\{\min[T^i_{(\Re_1 \circ \Re_2)}(h,g), T^i_{\Re_3}(g,t)]\}$$

$$= T^i_{(\Re_1 \circ \Re_2) \circ \Re_3}(h,t);$$

Note that $\Re_1 \circ \Re_2 \neq \Re_2 \circ \Re_1$, since the composition of two RNM relations $\Re_1, \Re_2$ exists, the composition of two RNM relations R_2,R_1 does not necessarily exist.

Phase 3: Development of Inverse Rough Neutrosophic Multisets Relation Theory

The inverse relation is formed by changing the element set of each of the ordered pairs in the given relation. Since we realize that relations are regularly indicated by relation and function, therefore, we prove that the RNM relation theory has an inverse relation.

Definition 7. Let Q be a non-empty set and X and Y be the rough neutrosophic multisets (RNM) in Q. $\Re \subseteq Q \times Q$ is RNM relation on Q×Q based on the X×Y. Then, we define $\Re^{-1} \subseteq Q \times Q$ as the RNM relation on Q×Q based on Y×X defined in Equation 7 as:

$$\mathfrak{R}^{-1} = \left\{ < (g,h), \left(T^{i}_{\mathfrak{R}^{-1}}(g,h)\right) \left(I^{i}_{\mathfrak{R}^{-1}}(g,h)\right), \left(F^{i}_{\mathfrak{R}^{-1}}(g,h)\right) >: (g,h) \in Q \times Q \right\} \quad (7)$$

where

$$T^{i}_{\mathfrak{R}^{-1}}(g,h) = T^{i}_{\mathfrak{R}}(h,g), I^{i}_{\mathfrak{R}^{-1}}(g,h) = I^{i}_{\mathfrak{R}}(h,g), F^{i}_{\mathfrak{R}^{-1}}(g,h) = F^{i}_{\mathfrak{R}}(h,g)$$

for all $(g,h) \in Q \times Q$ and $i = 1,2,\ldots,p$.

The relation R^{-1} is called the inverse rough neutrosophic multisets relation of R.

Proposition 7. $(\mathfrak{R}^{-1})^{-1} = \mathfrak{R}$.

Proof Since all the membership degree is true for all condition in definition 5; we only show the proving for truth membership sequence.

(1) i) $T^{i}_{(\mathfrak{R}^{-1})^{-1}}(h,g) = T^{i}_{\mathfrak{R}^{-1}}(g,h) = T^{i}_{\mathfrak{R}}(h,g) = 1$ **for all** $(h,g) \in \underline{X \times Y}$

ii) $T^{i}_{(\mathfrak{R}^{-1})^{-1}}(h,g) = T^{i}_{\mathfrak{R}^{-1}}(g,h) = T^{i}_{\mathfrak{R}}(h,g) = 0$ for all $(h,g) \in Q \times Q - \overline{X \times Y}$

iii) $0 < T^{i}_{(\mathfrak{R}^{-1})^{-1}}(h,g) = T^{i}_{\mathfrak{R}^{-1}}(g,h) = T^{i}_{\mathfrak{R}}(h,g) < 1$ for all $(h,g) \in \overline{X \times Y} - \underline{X \times Y}$

Thus $(\mathfrak{R}^{-1})^{-1} = \mathfrak{R}$.

Proposition 8. Let R_1,R_2 be two rough neutrosophic multisets relations on $Q\times Q$, based on $X\times Y$, $Y\times Z$, respectively. Then $(\mathfrak{R}_1 \circ \mathfrak{R}_2)^{-1} = {\mathfrak{R}_2}^{-1} \circ {\mathfrak{R}_1}^{-1}$.

Proof For all $h, g, \imath \in Q$, we have:

$T^{i}_{(\mathfrak{R}_1 \circ \mathfrak{R}_2)^{-1}}(\imath,h) = T^{i}_{\mathfrak{R}_1 \circ \mathfrak{R}_2}(h,\imath)$

$= \max_{g \in Q}\left\{\min\left[T^{i}_{\mathfrak{R}_1}(h,g), T^{i}_{\mathfrak{R}_2}(g,\imath)\right]\right\}$

$= \max_{g \in Q}\left\{\min\left[T^{i}_{(\mathfrak{R}_1)^{-1}}(g,h), T^{i}_{(\mathfrak{R}_2)^{-1}}(\imath,g)\right]\right\}$

$= \max_{g \in Q}\left\{\min\left[T^{i}_{(\mathfrak{R}_2)^{-1}}(\imath,g), T^{i}_{(\mathfrak{R}_1)^{-1}}(g,h)\right]\right\} = T^{i}_{(\mathfrak{R}_2)^{-1} \circ (\mathfrak{R}_1)^{-1}}(\imath,h);$

The proof is similar to indeterminacy function and falsity function.

Thus, it can be concluded that $(\mathfrak{R}_1 \circ \mathfrak{R}_2)^{-1} = {\mathfrak{R}_2}^{-1} \circ {\mathfrak{R}_1}^{-1}$.

Phase 4: Verifying Section by Illustrative Example of Rough Neutrosophic Multisets Relation Theory

In this section, the relation between three types of cars represented by the popularity of each car depending on expert opinion was studied. The definition of rough neutrosophic multiset (RNM) relation theory (definition 5) was used to get the best solution of the relation analysis. The flowchart of the algorithm used in solving RNM relation is shown in Figure 1 and an algorithm for each step is presented in Table 2.

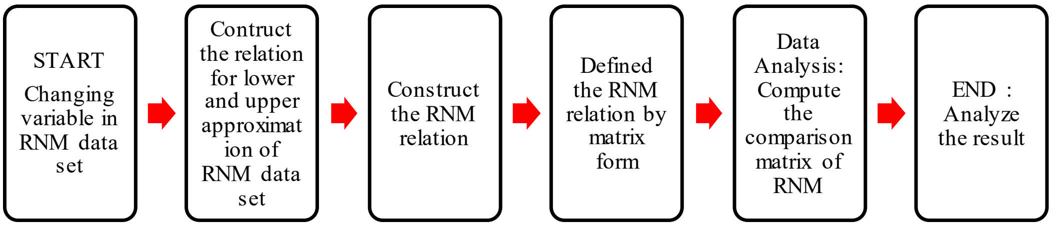


Figure 1. The flowchart of an algorithm in solving RNM relation

Table 2

An algorithm of steps to defined rough neutrosophic multisets relation.

Step		Algorithm
Step 1	:	Changing the variable in RNM data set: Compute lower and upper approximation values for rough neutrosophic multisets ($\underline{\mathcal{R}}_Q(X)$, $\underline{\mathcal{R}}_Q(Y)$, $\overline{\mathcal{R}}_Q(X)$, $\overline{\mathcal{R}}_Q(Y)$ by definition 3.
Step 2	:	Construct the relation of $\underline{X \times Y} = \underline{\mathcal{R}}_Q(X) \times \underline{\mathcal{R}}_Q(Y)$, relation of $\overline{X \times Y} = \overline{\mathcal{R}}_Q(X) \times \overline{\mathcal{R}}_Q(Y)$, and relation of $Q \times Q$. Note that $T^i_{\Re}(h, g) = 1$, $I^i_{\Re}(h, g) = 0$ and $F^i_{\Re}(h, g) = 0$ for all $(h, g) \in \underline{X \times Y}$ by definition 4 and 5.
Step 3	:	Construct a rough neutrosophic multisets relation $\Re$. Note that, $T^i_{\Re}(h, g) = 0$, $I^i_{\Re}(h, g) = 1$ and $F^i_{\Re}(h, g) = 1$ for all $(h, g) \in Q \times Q - \overline{X \times Y}$ by definition 5.
Step 4	:	Defined the RNM relation by matrix form: (i) Compute the unknown possible values by considering the condition of $0 < T^i_{\Re}(h, g) < 1$, for all $(h, g) \in \overline{X \times Y} - \underline{X \times Y}$ and neutrosophic multi relation of $T^i_{\Re}(h, g) \leq T^i_{X \times Y}(h, g)\ \forall\ (h, g) \in Q \times Q$. (ii) Defined $\Re \subseteq Q \times Q$ as a rough neutrosophic multisets relation on $Q \times Q$ based on the $X \times Y$ by a matrix form. We can have different values for unknown value based on different cases.
Step 5	:	Data Analysis: Compute the comparison matrix of RNM
Step 6	:	Analysis of the result

Consider the following example

Example 1. Assume that Q={1,2,3} be a universal set of types of car available for purchase in shop Z. R_Q be an equivalent relation on Q as hR_Q g if and only if h,g indicated the same popularities of car. hR_Q g is defined by R_Q={{1,3},{2}}. This relation justified that the car types 1 and 3 are in the same contingent and type 2 is in different contingent, but all the cars are still in the same universe. All the cars effected the profit of shop Z. Now, we try to get the opinion from two independent experts about the level of popularities of the car whether the car comprised "high", a level of indeterminacy with respect to the experts which was "not popular at all" and whether they felt that the car comprised "low level". Two – phases of the meeting were done in order to get the best result and the data was represented in neutrosophic multisets number. By satisfying all the condition in definition 3, we define the relation of rough neutrosophic multisets R on popularities of cars $Q\times Q$ based on experts' opinion $X\times Y$ as follows:

X=((1,0.3),(0.4,0.7),(0.6,0.8))/1+((0.5,0.7),(0.1,0.3),(0.4,0.5))/2+ ((1,0.6),(0.4,0.5),(0.6,0.7))/3,

Y=((0.4,0.6),(0.3,0.5),(0.1,0.7))/1+((0.5,0.4),(0.1,0.7),(0.3,0.8))/2+((1,0.7),(0.2,0.5),(0.1,0.7))/3

By following the algorithm presented (step 1 – 4), here we define a rough neutrosophic multisets relation R by a matrix.

We defined $\mathfrak{R} \subseteq Q \times Q$ as a rough neutrosophic multisets relation on $Q\times Q$ based on the X×Y by a matrix form:

$$M(\mathfrak{R}) = \begin{bmatrix} (0.4,0),(1,1),(1,1) & (0,0),(1,1),(1,1) & (0.6,0),(1,1),(1,1) \\ (0,0),(1,1),(1,1) & (0,0),(1,1),(1,1) & (0,0),(1,1),(1,1) \\ (0.5,0),(1,1),(1,1) & (0,0),(1,1),(1,1) & (0.9,0),(1,1),(1,1) \end{bmatrix}$$

Consider the RNM relation M(R) in example 1; then the inverse RNM relation R^{-1} of RNM relation R by using matrix is defined as:

Next, we compute the comparison matrix using the formula $D_{\mathfrak{R}}^i = T_{\mathfrak{R}}^i + I_{\mathfrak{R}}^i - F_{\mathfrak{R}}^i$ for all i, and select the maximum value for the comparison table. The result is shown in Table 3.

Table 3

Comparison matrix of rough neutrosophic multi relation R.

$\mathfrak{R}$	1	2	3
1	(**0.4**, 0.0)	(**0.0**, 0.0)	(**0.6**, 0.0)
2	(**0.0**, 0.0)	(**0.0**, 0.0)	(**0.0**, 0.0)
3	(**0.5**, 0.0)	(**0.0**, 0.0)	(**0.9**, 0.0)

By using row-sum and column-sum, the score for popularities of the car was computed as shown in Table 4.

RESULT AND DISCUSSION

The expert's opinion to decide the relationship between types of car and their popularities was successfully determined by rough neutrosophic multisets theory. According to the finding represented in Table 4, the experts agreed that the car type 1 was highly popular compared to car type 2 and type 3. This result was accurate since we had considered all uncertainty information level which level of popularities of the car was whether the car

Table 4
Popularity score for three types of car

	Row sum	Column sum	Popularity Score
1	1.0	0.9	**0.1**
2	0.0	0	0.0
3	1.3	1.5	(0.2)

comprised of "high", a level of indeterminacy with respect to the experts which is "not popular at all" and whether they feel that the car comprised of "low level". The repeated phase which was two – phases of the meeting was done in order to get the best result. Compared to others mathematical set data for uncertainty information such as fuzzy set, intuitionistic fuzzy set, neutrosophic set and rough neutrosophic set, our set data was more applicable, and it was already discussed in the literature. Therefore, the RNM relation theory could be used to represent the relation between the specified uncertainty information within the same universe.

The rough neutrosophic multisets (RNM) relation is a relation on neutrosophic multisets (NM), so we can consider it to be RNM relation over the universe. The RNM relation follows the condition of relation on NM, which is

Therefore, over universe Q, the RNM relation generalized relation for the rough neutrosophic set (RNS), rough intuitionistic fuzzy set (RIFS) and rough fuzzy set (RFS).

$T^i_{\mathfrak{R}}(h,g) \leq T^i_{X\times Y}(h,g), I^i_{\mathfrak{R}}(h,g) \geq I^i_{X\times Y}(h,g), F^i_{\mathfrak{R}}(h,g) \geq F^i_{X\times Y}(h,g)$ for all $(h,g) \in Q \times Q$, and

$0 \leq T^i_{\mathfrak{R}}(h,g) + I^i_{\mathfrak{R}}(h,g) + F^i_{\mathfrak{R}}(h,g) \leq 3.$

All the properties are true for RNS, RIFS and RFS.

1. The relation for RNS over universe Q is obtained when i=1 for all element T,I,F in definition 5 for Eq. (1),

R={<(h,g),(T_R (h,g))(I_R (h,g)),(F_R (h,g))>:(h,g) ∈ $Q\times Q$}.

2. The relation for RIFS over universe Q is obtained when i=1 for element T and F and condition (2) in definition 5 for Eq. (1) is omitted,

R={<(h,g),(T_R (h,g)),(F_R (h,g))>:(h,g) ∈ $Q\times Q$}.

3. The relation for RFS over universe Q is obtained when i=1 for element T and condition (2) and (3) in definition 5 is omitted, R={<(h,g),(T_R (h,g))>:(h,g) ∈ $Q\times Q$}.

CONCLUSION

Nowadays, there is much mathematical formulation developed to overcome the problem related to uncertainty information. Since the real data in the form of classical data is not suitable to represent the solution of uncertainty data, Zadeh (1965) introduced the fuzzy set data to overcome the inconsistent information in uncertainty problem. Then, Atanassov (1986) overcome the situation of inconsistent and incomplete information in the uncertainty problem. Next, Smarandache (1998) solved the issue of indeterminate information in uncertainty problem. The situation changed from single problem to repeated problem, and then a multiset approach was introduced for each of the set theory. All the fundamental theory approach has the same research direction to overcome the problem in uncertainty information that consists of inconsistent, incomplete and indeterminate information.

A consequence of that, this paper presents a new notion for uncertainty information by introduced rough neutrosophic multisets (RNM) relation into the universe. As the development of the RNM relation in RNM theory, the relationship developing RNM relation has primary consequences for both theories and application which is rough set and single valued neutrosophic multisets.

The important properties for fundamental theory was presented in this paper. We had computed the operation of min, max, the composition of two RNM and inverse relation for RNM. An algorithm with an illustrative example of relation analysis of popularities of cars used by different expert opinions is presented to help other researchers better understand the RNM relation theory procedure. The result shows a significant relationship for cars used based on popularity. The discussion has indicated that RNM relation also allows for the generalization of the RNS relation, RIFS relation, and RFS relation. This generalization helps other researchers to elaborate more on relations involving uncertainty information.

RNM relation is studied to verify the value of the interactions between variable use in neutrosophic multiset with rough approximation, such as the relationship between patient and symptoms in the different time taken to diagnose the right patient disease. RNM relation is a useful tool in dealing with multiple phases of inconsistent, incomplete and indeterminate information.

In future work, it would be meaningful to explore further RNM theory properties such as distance and similarity measure and their applications in solving uncertainty information.

ACKNOWLEDGEMENT

The authors would like to acknowledge Universiti Teknologi MARA and Ministry of Higher Education, Malaysia, for their financial support. Also, thanks go to the anonymous reviewers and the editor for their generous ideas and comments towards improving the clarity of this paper.

REFERENCES

Ali, M., & Smarandache, F. (2017). Complex neutrosophic set. *Neural Computing and Applications, 28*(7), 1817-1834.

Alias, S., Mohamad, D., & Shuib, A. (2017). Rough neutrosophic multisets. *Neutrosophic Sets and Systems, 16,* 80-88.

Alkhazaleh, S. (2016). Time-neutrosophic soft set and its applications. *Journal of Intelligent and Fuzzy Systems, 30*(2), 1087-1098.

Atanassov, K. T. (1986). Intuitionistic fuzzy sets. *Fuzzy Sets and Systems, 20*(1), 87-96.

Broumi, S., & Smarandache, F. (2014). Interval neutrosophic rough set. *Neutrosophic Sets and Systems, 7,* 23-31.

Broumi, S., Smarandache, F., & Dhar, M. (2014). Rough neutrosophic sets. *Italian Journal of Pure and Applied Mathematics, 32,* 493-502.

Guo, Z. L., Liu, Y. L., & Yang, H. L. (2017). A novel rough set model in generalized single valued neutrosophic approximation spaces and its application. *Symmetry, 9*(7), 119-137.

Mandal, D. (2015). Comparative study of intuitionistic and generalized neutrosophic soft sets. International *Journal of Mathematical, Computational, Natural and Physical Engineering, 9*(2), 111-114.

Molodtsov, D. (1999). Soft set theory - first results. *An International Journal of Computers and Mathematics with Applications, 37,* 19-31.

Mondal, K., & Pramanik, S. (2015). Rough neutrosophic multi-attribute decision-making based on rough accuracy score function. *Neutrosophic Sets and Systems, 7,* 8-17.

Mondal, K., Pramanik, S., & Smarandache, F. (2016). Multi-attribute decision making based on rough neutrosophic variational coefficient similarity measure. *Neutrosophic Sets and Systems, 13,* 3-17.

Nguyen, X. T., Nguyen, V. D., & Nguyen, D. D. (2014). Rough fuzzy relation on two universal sets. *International Journal of Intelligent Systems and Applications, 6(*4), 49-55.

Pawlak, Z. (1982). Rough sets. *International Journal of Computer and Information Sciences, 11*(5), 341-356.

Pramanik, S., & Mondal, K. (2015). Cosine similarity measure of rough neutrosophic sets and its application in medical diagnosis. *Global Journal of Advanced Research, 2*(1), 212-220.

Shinoj, T. K., & John, S. J. (2012). Intuitionistic fuzzy multisets and its application in medical diagnosis. *World Academy of Science, Engineering and Technology, 6*(1), 1418-1421.

Smarandache, F. (1998). *Neutrosophy, neutrosophic probability, set, and logic* (6th Ed.). Michigan, USA: ProQuest Information & Learning.

Smarandache, F. (2013). n-Valued refined neutrosophic logic and its applications in physics. *Progress in Physics, 4,* 143-146.

Yager, R. R. (1986). On the theory of bags. *International Journal of General Systems, 13*(1), 23-37.

Yang, H., Guo, Z., & Liao, X. (2016). On single valued neutrosophic relations. *Journal of Intelligent and Fuzzy Systems, 30*(2), 1-20.

Yang, H. L., Zhang, C. L., Guo, Z. L., Liu, Y. L., & Liao, X. (2017). A hybrid model of single valued neutrosophic sets and rough sets: single valued neutrosophic rough set model. *Soft Computing, 21*(21), 6253-6267.

Ye, S., & Ye, J. (2014). Dice similarity measure between single valued neutrosophic multisets and its application in medical diagnosis. *Neutrosophic Sets and Systems, 6*, 48-53.

Zadeh, L. A. (1965). Fuzzy sets. *Information and Control, 8*(3), 338-353.

Pertanika J. Sci. & Technol. 27 (4): 2373 - 2379 (2019)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Short Commucation

Characterization of Model on Zoning Process

Ompal Singh

Department of Mathematics, SRM Institute of Science and Technology, Modinagar Ghaziabad, Uttar Pradesah 201204, India

ABSTRACT

This paper represents the process by which a network is partitioned into smaller network each of which further delegated with a smaller network with a certain degree of autonomy in terms of resource allocation and operation. The term "Autonomy" implies that once the guiding policy is articulated and the resource allocation is decided upon, local management may enjoy some freedom in local, short-term decisions such as dispatching repositioning, budget planning and manning. The implications of zoning prevail over a long period, once a wide network is partitioned into sub network, each sub network will likely be treated as almost an independent network in terms of its "rights" to possess and to operate resources.

Keywords: Autonomy, delegated, network, partition, resources

ARTICLE INFO

Article history:
Received: 3 May 2018
Accepted: 21 June 2019
Published: 21 October 2019

E-mail address:
ompalphd@gmail.com

INTRODUCTION

One major advantage of zoning is that it facilitates the modeling and the solving of the local network policy problems. In most cases it consumes less time and effort than an attempt to use a global model. In the context of providing services, particularly in the public sector, the concept of equity asserts that the entire population of potential clients be treated as equally as possible in terms of the quality of service they get. Apply the equity criterion to a service network implies that the performance measures by which

ISSN: 0128-7680
e-ISSN: 2231-8526

the quality of service is evaluated be more or less equal in each sub zone (Ayeni, 1976). Practical realization of this criterion could be accomplished by partitioning the network into sub network that are more or less equal in the proportion of demand they generate. The sample network G exhibited in Figure 1.

Network G consists of 9 nodes and 16 links. The nodes have been numbered arbitrarily from 1 to 9. The Figure 1 near the links designates the length of the links. We will denote them l (i, j), where i and j are node numbers. The fractions near the nodes indicate the proportions of the total demand generated in the particular nodes. These are denoted by P_j, j = 1,……., 9. Note that $\sum_{j=1}^{I} P_{j=1}$. We will certainly not recommend that node 2 and 9 constitute G while all the rest of the nodes are assigned to G since such partitioning will load 81% of the total demand G. Rather; we will try to mark nodes such that their accumulated demand will be close to 50%.

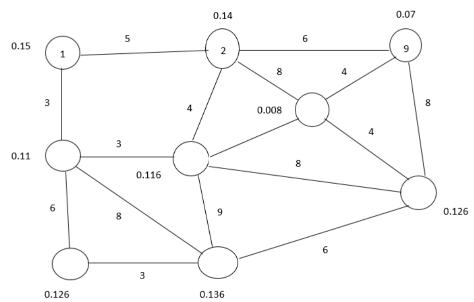


Figure 1. A sample network G

Contiguity

A basic principle in zoning is contiguity. A sub network is contiguous if it is possible to travel from every node in the sub network to every other node in it without crossing another sub network. In other words, there should be at least one path between any two nodes of a sub network such that a server is able to travel between the two nodes on that path without having to go through another sub network (Dutta & Kumar, 2015). This is not to say that this path is necessarily the shortest one. It may very well be that the shortest path (Beckman, 1981) will be across another network, but there is at least one more path that is under the sovereignty of the said network, therefore contiguity is satisfied. The major reason is that it allows the sub network management to move its servers along the network without having to get permission or to coordinate the move with foreign authorities. Thus, dispatching, patrolling, and repositioning policies can be devised independently. One possible way to contiguity is by constructing a square matrix whose elements are binary, namely, zero or ones (Table 1). The rows and the columns correspond to the nodes of the network.

Compactness

An intuitive interpretation of the notion compactness is that the edges of a zone are not too remote from each other. In partitioning a planar area (rather than a network) compactness can be measured by any of three measurements:

(a) Resemblance of the zone to a square.
(b) Resemblance of the zone to a circle.
(c) Reasonable" distance of the population from the center of the zone.

In network partitioning, managements related to a planar area topology do not adhere to the notion of a network.

A certain degree of proximity is maintained among the nodes of a zone. This can be obtained by imposing a length constraint on the shortest distance between any two nodes (Berman & Odoni, 1982) that are considered candidates for belonging to the same zone. Table 2 displays the shortest distance between any two nodes of network G. Based on Table 2 we can impose an arbitrary length beyond which two nodes cannot be part of the same zone. Suppose the arbitrary limit is 10. In Figure 1 we mark by 1 all the element in Table 2 that are less than or equal to 10, and by 0 all the elements that are greater than 10. Figure 1 constraints the zoning decision, namely, upon examining a node that is a candidate to be selected to a certain zone, we will observe with the compactness criterion. Such a table is called an exclusion matrix. The values of its element depend, of course, on the arbitrary value of the length constraint.

Table 1

A contiguity matrix

Node \ Node	1	2	3	4	5	6	7	8	9
1	**-**	**1**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
2	**1**	**-**	**0**	**1**	**1**	**0**	**0**	**0**	**1**
3	**1**	**0**	**-**	**1**	**0**	**1**	**1**	**0**	**0**
4	**0**	**1**	**1**	**-**	**1**	**0**	**1**	**1**	**0**
5	**0**	**1**	**0**	**1**	**-**	**0**	**0**	**1**	**1**
6	**0**	**0**	**1**	**0**	**0**	**-**	**1**	**0**	**0**
7	**0**	**0**	**1**	**1**	**0**	**1**	**-**	**1**	**0**
8	**0**	**0**	**0**	**1**	**1**	**0**	**1**	**-**	**1**
9	**0**	**1**	**0**	**0**	**1**	**0**	**0**	**0**	**-**

Table 2

Shortest distance in G

From					**To**				
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
1	**0**	**5**	**3**	**6**	**10**	**9**	**11**	**14**	**11**
2	**5**	**0**	**7**	**4**	**8**	**13**	**13**	**12**	**6**
3	**3**	**7**	**0**	**3**	**7**	**6**	**8**	**11**	**11**
4	**6**	**4**	**3**	**0**	**4**	**9**	**9**	**8**	**8**
5	**10**	**8**	**7**	**4**	**0**	**13**	**10**	**4**	**4**
6	**9**	**13**	**6**	**9**	**13**	**0**	**3**	**9**	**17**
7	**11**	**13**	**8**	**9**	**10**	**3**	**0**	**6**	**14**
8	**14**	**12**	**11**	**8**	**4**	**9**	**6**	**0**	**8**
9	**11**	**6**	**11**	**8**	**4**	**17**	**14**	**8**	**0**

Enclaves

During the process of zoning we have to make sure that we do not create enclave (Sarkar et al., 2015). An enclave is a node, or a subset of nodes that cannot constitute an independent zone because of the equity criterion. On the other hand, the node cannot be connected to other "free" nodes for no contiguity reasons. Thus, they might remain "Orphans" if the zoning process proceeds without being interrupted.

ADDITIONAL CRITERIA

There could be some additional terms that a network planner would be required to accede to under certain circumstances. Take for instance administrative boundaries. Another criterion is related to the characteristics of the region being partitioned. It is therefore recommended to account for the characteristics (Conway et al., 1996) of the region before a "mechanical" zoning process is executed. When these or similar criteria are being examined, one has to distinguish between mandatory requirements and optimal requirement. A mandatory requirement possesses a constraint that must be followed. In a way, a mandatory requirement can sometimes facilitate the computational complexity of a zoning algorithm because it usually splits the problem into a number of smaller problems, each of which can handled more easily. An optional requirement is unlikely to facilitate the solving process-it is more likely to complicate it (Mishra & Tripathy, 2012). The planner has to solve the constrained model as well as the unconstrained one is order to provide the decision maker with the "cost" of the additional requirement, cost in this respect is a decrease in performance. Nonetheless, with the fast advance of computing technology, running an algorithm for a number of times under varied constraints is usually not infeasible. We turn now to introducing a zoning selection algorithm.

An Algorithm for Zoning Selection

Zoning selection process has been applied mostly for area districting. An elaborate algorithm for such cases was provided by Garfinkel and Nemhauser (1970). When network zoning is considered, however, some of the guiding criteria have to be modified (Eilon et al.., 1971). For the notion of compactness is expressed in distance measurement rather than in area topology, the notion of contiguity is expressed by connectivity of nodes rather than by having common borders.

Let us try to partition network G in to four zones. We impose the following constraints:

(a) Equity: The 'ideal' demand generated in each zone would be 25% however, we allow for 2.5% deviation, namely, an acceptable zone may generate demand ranging from 22.5% to 27.5% of the total demand.

(b) Contiguity: Contiguity must be maintained for each zone.

Compactness

The shortest distance between any two nodes in a zone should not exceed 10 units of time. Here enclaves must be avoided during the application of the algorithm. Suppose these are the only restrictions imposed on the zoning process. The partitioning process consists of two major phases (EI-Shaieb, 1973). Phase I, determines all the possible zones that comply with the requirement listed above. Since this phase identifies all the possible zones, upon completion of Phase I, we may very well face redundancy in node coverage, that is, a certain node might belong to more than one candidate zone. Phase II, given the required number of zones, determines the node partitioning according to a suitable objective function (Hallpern & Maimon, 2011). It therefore eliminates some of the possible zones in order to obtain not only an exhaustive but also a mutually exclusive partitioning. The way Phase I works is by selecting an arbitrary node and trying to form all the feasible zones that include that node while watching not to violate any constraint.

The nodes are selected in an ascending order according to their arbitrary numbers from 1 to 9. In this manner we make sure that nodes are not being overlooked; also, when a certain node is selected, the process has to examine only subsequent nodes (in terms of their serial numbers) because it is guaranteed that preceding nodes have been examined previously (Garfinkel and Nemhauser, 1970). Let us start, then, with node 1. It is linked to node 2, together they accumulate 27% of the total demand. They do not violate compactness, nor do they enclave any node, thus {1, 2} constitute a feasible zone (Hall, 2009). We cannot add any more node [1, 2} since any additional node that is linked either to 1 or to 2 will push the demand beyond the tolerated limit, which is 27.5%. By similar arguments, nodes {1, 3} form a feasible zone that cannot be further augmented. Let us turn now to node 2. We do not have to examine the combination of 2 and 1 becomes this has already been covered. New feasible zones are, therefore, {2, 4} and {2, 5, 9}.

Table 3 summarizes (Goldman, 1971) the final results of phase I, the set of feasible zones .In column (a) we have numbered the zone and column (b) designates the number of the zone and column (c) displays the total demand of a zone. And column (d) calculates the amount of deviation of the demand relative to maximum tolerated deviation (2.5%).

$$\frac{|29 - 25.2|}{2.5} = \frac{3.8}{2.5} = 1.52$$

Column (e) displays the largest shortest distance, namely, the shortest distance between the most remote nodes within a zone. The table is divided into sections. Each section is associated with another “root node”, namely, a node from the search for feasible zones begins. The “root node” determines the section number in column (f).

Table 3

List of feasible zones

Zone No.(a)	Nodes (b)	Demand(%),(c)	Relative deviation (d)	Largest shortest distance (e)	Section list (f)
1	1,2	29	0.8	5	1
2	1,3	26	0.4	3	-
3	2,4	25.6	0.2	4	2
4	2,5,9	29	0.4	8	-
5	3,6	23.6	1.0	6	3
6	4,5,9	26.6	0.2	8	4
7	4,7	25.2	00	9	-
8	4,8	24.2	0.4	8	-
9	5,8,9	27.6	0.2	8	5
10	6,7	26.2	0.4	3	6
11	7,8	26.2	0.4	6	7

CONCLUSION

Eventually, both optimizations of equity and compactness have provided the same partitioning. This, however, is not necessarily the case a more complex problem is encountered. Generally speaking, by comparing the results of a number of optimization processes, one can learn the "price" paid in one criterion in order to optimize another one. The zoning process that is presented above was performed mainly by observation.

ACKNOWLEDGEMENT

The authors would like to thank the anonymous reviewers for their critical comments and valuable suggestions. The general and the technical remarks of the editor and reviewers were of great help for improving the quality and presentation of the paper.

REFERENCES

Ayeni, M. A. O. (1976). A predictive model of urban stock and activity: 2. Empirical development. *Environment and Planning A, 8*(1), 59-77.

Beckman, M. J. (1981). Continuous Model of transportation and location. *International Journal of Network System, 4*(5), 619-630.

Berman, O., & Odoni, A. R. (1982). Locating mobile servers on a network with Markovian properties. *Networks, 12*(1), 73-86.

Conway, R. W., Maxwen, W. O., & Miller, L.W. (1996). Theory of scheduling. *International Journal of Network System, 4*(5), 312-325.

Dutta, D., & Kumar, P. (2015). A partial backlogging inventory model for deteriorating items with time varying demand and holding cost. *International Journal of Mathematics in Operational Research, 7*(3), 281-296.

Eilon, S., Watson-Gandy, C. D. T., & Christofides, N. (1971). *Distribution Management: Mathematical Modeling and Practical analysis*. London: Griffin.

El-Shaieb, A. M. (1973). A new algorithm for locating sources among destinations. *Management Science, 20*(2), 221-231.

Garfinkel, R. S., & Nemhauser, G. L. (1970). Optimal political districting by implicit enumeration techniques. *Management Science 16*(8), 495- 508.

Goldman, A. J. (1971). Optimal centre location in simple networks. *Transportation Science, 5*(2), 212-221.

Hall, R. W. (2009). Direct versus Terminal Freight routing on a network with concave costs. *International Journal of Network System, 4*(5), 287- 298.

Hallpern, J., & Maimon, O. (2011). Algorithms for the m-centre problems a survey. *European Journal of Operation Research, 10*(3), 195-203.

Mishra, U., & Tripathy, C. K. (2012). An EOQ model for time dependent Weibull deterioration with linear demand and shortages. *Scientific Journal of Logistics, 8*(2), 123-136.

Sarkar, B., Saren, S., & Cardenas-Barron, L. E. (2015). An inventory model with trade credit policy and variable deterioration for fixed lifetime products. *Annals of Operations Research, 229*(1), 677-702.

Pertanika J. Sci. & Technol. 27 (4): 2381 - 2398 (2019)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Two Point Block Multistep Methods with Trigonometric–Fitting for Solving Oscillatory Problems

Aini Fadhlina Mansor[1], Fudziah Ismail[1, 2]* and Norazak Senu [1, 2]

[1]*Institute for Mathematical Research, Universiti Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia*
[2]*Department of Mathematics, Faculty of Science, Universiti Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia*

ABSTRACT

In this paper, we present the absolute stability of the existing 2-point implicit block multistep step methods of step number $k = 3$ and $k = 5$ and solving special second order ordinary differential equations (ODEs). The methods are then trigonometrically fitted so that they are suitable for solving highly oscillatory problems arising from the special second order ODEs. Their explicit counterparts are also trigonometrically fitted so that in the implementation the methods can act as a predictor-corrector pairs. The numerical results based on the integration over a large interval are given to show the performance of the proposed methods. From the numerical results we can conclude that the new trigonometrically-fitted methods are superior in terms of accuracy and execution time, compared to the existing methods in the scientific literature when used for solving problems which are oscillatory in nature.

Keywords: Block method, multistep method, oscillatory problems, special second order ODEs, trigonometrically fitted

INTRODUCTION

In this research, we are concerned with the numerical methods for solving special second order ordinary differential equation (ODE) of the form as follows:

$$y'' = f(x, y),\ y(x_0) = y_0, y'(x_0) = y'_0, \qquad [1]$$

General second order ODE can be written as $y'' = f(x, y, y')$, special second order ODE does not depend on the derivative of the solution. This type of ordinary differential equations often appear in many scientific areas such as mechanics,

ARTICLE INFO

Article history:
Received: 06 December 2018
Accepted: 04 June 2019
Published: 21 October 2019

E-mail addresses:
aini_virgo92@yahoo.com (Aini Fadhlina Mansor)
fudziah_i@yahoo.com.my (Fudziah Ismail)
norazak@upm.edu.my (Norazak Senu)
* Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

astrophysics, quantum chemistry and electronics. Further details can be seen in Fang and Wu (2008).

The most common technique to numerically solve the second order problem is by reducing it to a system of first order ODEs. However it is more efficient if [1] can be solved directly without converting it to a system of first order ODEs. Such methods for directly solving the special second order ODEs are direct multistep method, Runge-Kutta-Nyström (RKN) method and hybrid method.

All the methods mentioned above, approximate the solution of the equation at only one point at a time step. Hence, to increase the efficiency of the numerical methods for solving ODEs, recently there is a lot of research has been done on block methods. Block multistep methods calculate the solutions of the ODEs at more than one point at a time, for example 2 point block method calculates the solution at two points concurrently, hence less execution time is needed to solve the ODEs. Fatunla (1995) constructed block methods for solving special second order ordinary differential equations. Then, Akinfewa et al. (2015) proposed a family of continuous third derivative block methods for numerical integration of first order system of ODEs. Ramos et al. (2015) developed an optimized two-step hybrid block method for solving general second order ODEs.

Quite often the solution of [1] exhibits a pronounced oscillatory character. Oscillatory problems are usually harder to solve than the non oscillatory problems. To obtain a more efficient process for solving oscillatory problems, numerical methods are constructed by taking into account the nature of the problem. This results in methods in which the coefficients depend on the frequency of the problem to be solved. Some important classes of the numerical methods are exponentially fitted, trigonometrically fitted or phase-fitted methods.

Simos (2003) in his work, had proposed exponentially-fitted and trigonometrically-fitted symmetric linear multistep methods for solving orbital problems. Then, Fang and Wu (2008) in their research, had proposed trigonometrically fitted explicit Numerov-type method. Phase fitted Runge-Kutta-Nystrom method had been studied by Papadopoulos et al. (2008), Ahmad et al. (2016), extended the work by developing trigonometrically-fitted hybrid method for solving oscillatory delay differential equations.

In the quest for methods that best approximate the solution of [1] which are oscillatory in nature, in this paper, we developed an efficient block multistep methods by trigonometrically-fitted the methods, in which to the best of our knowledge is the first work on trigonometrically fitted block multistep methods for solving highly oscillatory problems. The coefficients of the method depend on the frequency of the problem to be solved, hence the frequency of the problems must be priory known.

The research is based on the 2-point and 3-point explicit and implicit block multistep methods which have been derived in Mansor et al. (2017). First, the stability aspect of the 2-point and 3-point implicit block multistep methods are investigated. Then, both the explicit and implicit block multistep methods of step number $k = 3$ and $k = 5$, are trigonometrically fitted. The methods are then implemented as a predictor-corrector pairs. Numerical results of the proposed methods in the form of efficiency curves for solving five oscillatory problems are then presented.

MATERIALS AND METHODS

Absolute Stability

In our previous work, see Mansor et al. (2017), we derived the 2-point and 3-point explicit and implicit block multistep methods of step number $k = 3$ and $k = 5$ with orders three and five respectively for the explicit methods and orders four and six for the implicit methods. Here we are going to investigate the stability of the implicit block methods.

2-point implicit Block Methods for $k = 3$:

The first and second point of the implicit block methods that have been derived in Mansor et al. (2017) for step number $k = 3$ are given as follows:

$$y_{n+1} = 2y_n - y_{n-1} + h^2\left(\frac{1}{12}f_{n+1} + \frac{5}{6}f_n + \frac{1}{12}f_{n-1} + 0f_{n-2}\right), \quad [2]$$

$$y_{n+2} = 2y_n - y_{n-2} + h^2\left(0f_{n+2} + \frac{4}{3}f_{n+1} + \frac{4}{3}f_n + \frac{4}{3}f_{n-1}\right). \quad [3]$$

The implicit 2-point block methods can be presented in the matrix form as:

$$\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}\begin{bmatrix}y_{n+1}\\y_{n+2}\end{bmatrix} = \begin{bmatrix}-1 & 2\\0 & 2\end{bmatrix}\begin{bmatrix}y_{n-1}\\y_n\end{bmatrix} + \begin{bmatrix}0 & 0\\0 & -1\end{bmatrix}\begin{bmatrix}y_{n-3}\\y_{n-2}\end{bmatrix}$$
$$+h^2\left(\begin{bmatrix}\frac{1}{12} & 0\\ \frac{4}{3} & 0\end{bmatrix}\begin{bmatrix}f_{n+1}\\f_{n+2}\end{bmatrix} + \begin{bmatrix}\frac{1}{12} & \frac{5}{6}\\ \frac{4}{3} & \frac{4}{3}\end{bmatrix}\begin{bmatrix}f_{n-1}\\f_n\end{bmatrix}\right)$$

Using the test equation $y'' = \lambda y$ on the method, gives

$$\begin{bmatrix}1 - \frac{1}{12}\lambda h^2 & 0\\ -\frac{4}{3}\lambda h^2 & 1\end{bmatrix}\begin{bmatrix}y_{n+1}\\y_{n+2}\end{bmatrix} = \begin{bmatrix}-1 + \frac{1}{12}\lambda h^2 & 2 + \frac{5}{6}\lambda h^2\\ \frac{4}{3}\lambda h^2 & 2 + \frac{4}{3}\lambda h^2\end{bmatrix}\begin{bmatrix}y_{n-1}\\y_n\end{bmatrix} + \begin{bmatrix}0 & 0\\0 & -1\end{bmatrix}\begin{bmatrix}y_{n-3}\\y_{n-2}\end{bmatrix}.$$

Taking the determinant equals to zero, we have the stability polynomial $t^2\left(-\frac{1}{12}\hbar + 1\right) + t(-\hbar^2 - 4\hbar) + \left(-\hbar^2 - \frac{47}{12}\hbar - 1\right) = 0$

where $\hbar = \lambda h^2$.

Then, by solving the stability polynomial for values of $\hbar$ with $|t| \leq 1$, gives the absolute stability region of the method as shown in Figure 1, where the horizontal axis is the real part of $\hbar$ and the vertical axis is the imaginary part of $\hbar$.

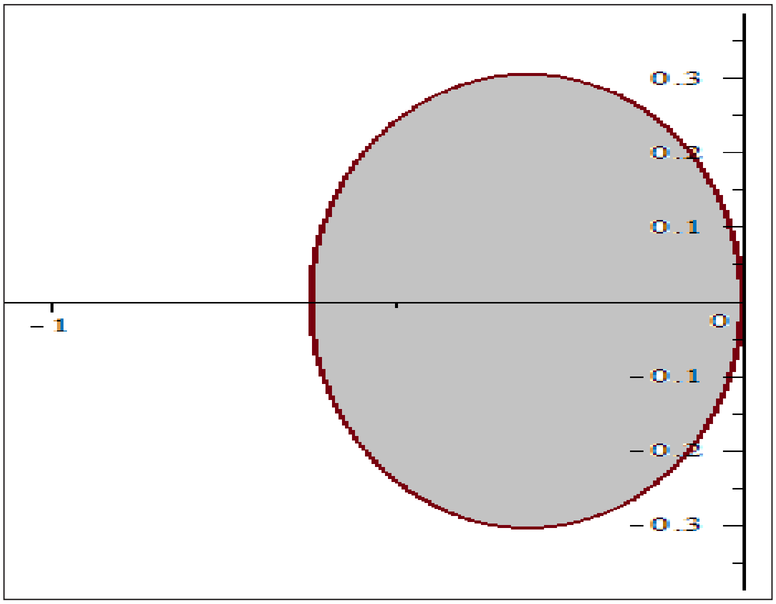


Figure 1. Stability region of the 2-point implicit block multistep method for $k = 3$

The 2-point implicit block multistep method for $k = 3$, has a small region of absolute stability, however it is still stable and can be used to solve the special second order ODEs.

2-point Block Methods for $k = 5$:

$$y_{n+1} = 2y_n - y_{n-1} + h^2\left(\frac{3}{40}f_{n+1} + \frac{209}{240}f_n + \frac{1}{60}f_{n-1} + \frac{7}{120}f_{n-2} - \frac{1}{40}f_{n-3} + \frac{1}{240}f_{n-4}\right)\Big] \quad [4]$$

$$y_{n+2} = 2y_n - y_{n-2} + h^2\left(\frac{1}{15}f_{n+2} + \frac{16}{15}f_{n+1} + \frac{26}{15}f_n + \frac{16}{15}f_{n-1} + \frac{1}{15}f_{n-2} + 0f_{n-3}\right) \quad [5]$$

The implicit 2-point block methods can be presented in the matrix form as:

$$\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}\begin{bmatrix}y_{n+1}\\y_{n+2}\end{bmatrix} = \begin{bmatrix}-1 & 2\\0 & 2\end{bmatrix}\begin{bmatrix}y_{n-1}\\y_n\end{bmatrix} + \begin{bmatrix}0 & 0\\0 & -1\end{bmatrix}\begin{bmatrix}y_{n-3}\\y_{n-2}\end{bmatrix} + h^2\left(\begin{bmatrix}\frac{3}{40} & 0\\\frac{16}{15} & \frac{1}{15}\end{bmatrix}\begin{bmatrix}f_{n+1}\\f_{n+2}\end{bmatrix} + \begin{bmatrix}\frac{1}{60} & \frac{209}{240}\\\frac{16}{15} & \frac{26}{15}\end{bmatrix}\begin{bmatrix}f_{n-1}\\f_n\end{bmatrix}\right.$$
$$\left. + \begin{bmatrix}-\frac{1}{40} & \frac{7}{120}\\0 & \frac{1}{15}\end{bmatrix}\begin{bmatrix}f_{n-3}\\f_{n-2}\end{bmatrix}\right).$$

Substituting the test equation $y'' = \lambda y$ into the method, gives,

$$\begin{bmatrix} 1-\frac{3}{40}\lambda h^2 & 0 \\ -\frac{16}{15}\lambda h^2 & 1-\frac{1}{15}\lambda h^2 \end{bmatrix}\begin{bmatrix} y_{n+1} \\ y_{n+2} \end{bmatrix} = \begin{bmatrix} -1+\frac{1}{60}\lambda h^2 & 2+\frac{209}{240}\lambda h^2 \\ \frac{16}{15}\lambda h^2 & 2+\frac{26}{15}\lambda h^2 \end{bmatrix}\begin{bmatrix} y_{n-1} \\ y_n \end{bmatrix}$$
$$+\begin{bmatrix} -\frac{1}{40}\lambda h^2 & \frac{7}{120}\lambda h^2 \\ 0 & -1+\frac{1}{15}\lambda h^2 \end{bmatrix}\begin{bmatrix} y_{n-3} \\ y_{n-2} \end{bmatrix}$$
$$+\begin{bmatrix} 0 & \frac{1}{240}\lambda h^2 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} y_{n-5} \\ y_{n-4} \end{bmatrix}.$$

Taking the determinant equals to zero, we have the stability polynomial

$$t^2\left(\frac{1}{200}\hbar^2-\frac{17}{120}\hbar+1\right)+t\left(-\frac{31}{36}\hbar^2-\frac{47}{12}\hbar\right)+\left(-\frac{1819}{1800}\hbar^2-\frac{473}{120}\hbar-1\right)=0.$$

Solving the stability polynomial for values of $\hbar$ with $|t| \leq 1$, we obtained the absolute stability region of the method as shown in the Figure 2, where the horizontal axis is the real part of $\hbar$ and the vertical axis is the imaginary part of $\hbar$.

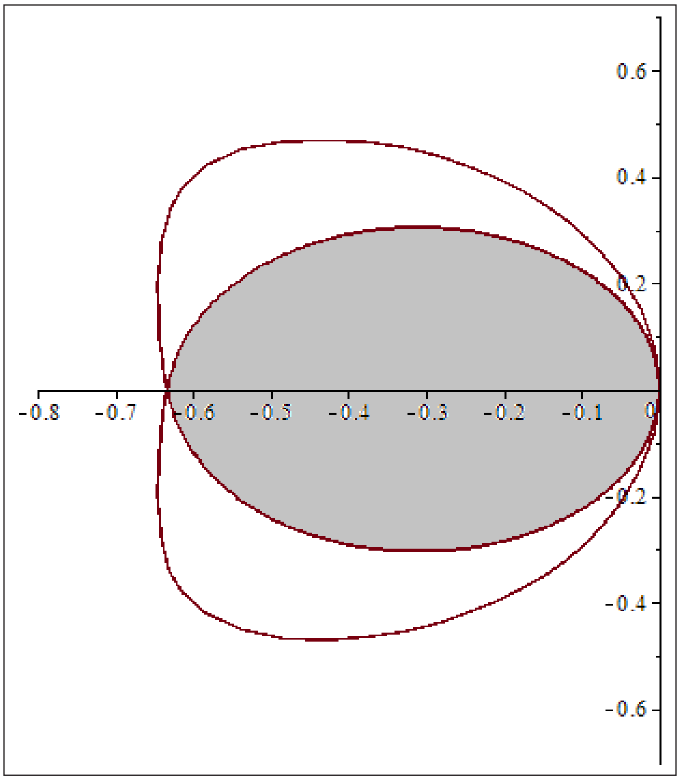


Figure 2. Stability region of the 2-point implicit block multistep method for $k = 5$

Hence we can conclude that both the k = 3 and k = 5 step implicit block multistep methods do have a substantial regions of absolute stability. In the next section, we will derive the trigonometrically ftted block methods.

Derivation of the Trigonomerically-Fitted Methods

Explicit block method for k = 3:

The first and second point of the explicit block methods that have been derived in Mansor et al. (2017) for step number k = 3 are given as follows:

$$y_{n+1} = 2y_n - y_{n-1} + h^2\left(\frac{13}{12}f_n - \frac{1}{6}f_{n-1} + \frac{1}{12}f_{n-2}\right) \quad [6]$$

$$y_{n+2} = 2y_n - y_{n-2} + h^2\left(\frac{16}{3}f_n - \frac{8}{3} - f_{n1} + \frac{4}{3}f_{n-2}\right) \quad [7]$$

In general form, the methods can be written as

$$y_{n+1} = 2y_n - y_{n-1} + h^2(p_2 f_n + p_1 f_{n-1} + p_0 f_{n-2}) \quad [8]$$

$$y_{n+2} = 2y_n - y_{n-2} + h^2(q_2 f_n + q_1 f_{n-1} + q_0 f_{n-2}) \quad [9]$$

Equation $y'' = f(t, y) = -\omega^2 y$ is the equation that most researchers used in the literature when they are dealing with oscillatory problems and trigonometric-fitting methods. For further details see Li et al. (2017). The method integrates exactly the differential equation whose solutions can be expressed as the linear combination of $\{\sin(\omega t), \cos(\omega t)\}$. Hence $y_n = \mathrm{co\,s}(\omega t_n)$, we have

$$y_n' = -\omega \sin(\omega t_n)$$
$$y_n'' = -\omega^2 \cos(\omega t_n)$$
$$y_n'' = f(t_n, y_n) = -\omega^2 y_n = f_n$$
$t_{n+1} = t_n + h$ and $t_{n-1} = t_n - h$.

Let $H = \omega h$ and taking $t_n = 0$, this is the same approach as in Fang and Wu (2008), thus we have

$y_n = \cos(0)$, $y_{n-1} = \cos(-H)$, and $y_{n+1} = \cos(H)$.

Substituting into [8] and [9], we obtain

$$2\cos(H) = 2 - H^2\big(p_2 + p_1\cos(H) + p_0\cos(2H)\big), \quad [10]$$

$$2\cos(2H) = 2 - H^2\big(q_2 + q_1\cos(H) + q_0\cos(2H)\big). \quad [11]$$

Then, letting $y_n = \sin(\omega t_n)$ and using the same technique, we obtain

$$H^2\big(p_1\sin(H)+p_0\sin(2H)\big)=0, \qquad [12]$$

$$H^2\big(q_1\sin(H)+q_0\sin(2H)\big)=0. \qquad [13]$$

There are six undetermined coefficients, they are $p_0, p_1, p_2, q_0, q_1, q_2$ and four equations to be solved. Letting p_0 and q_0 as free parameters where the values are obtained from the coefficients of the original methods ($p_0=\frac{1}{12}$, $q_0=\frac{4}{3}$) and solving equations [10] - [13] and rewriting in Taylor series expansion to avoid heavy cancellation in the implementation of the methods, we have

$$p_1=-\frac{1}{6}+\frac{1}{12}H^2-\frac{1}{144}H^4+\frac{1}{4320}H^6-\frac{1}{241920}H^8+\frac{1}{21772800}H^{10}-\frac{1}{2874009600}H^{12}+O(H^{14}),$$

$$p_2=\frac{13}{12}-\frac{1}{12}H^2+\frac{1}{360}H^4-\frac{1}{20160}H^6+\frac{1}{1814400}H^8-\frac{1}{239500800}H^{10}+O(H^{12}),$$

$$q_1=-\frac{8}{3}+\frac{4}{3}H^2-\frac{1}{9}H^4+\frac{1}{270}H^6-\frac{1}{15120}H^8+\frac{1}{1360800}H^{10}-\frac{1}{179625600}H^{12}+O(H^{14}),$$

$$q_2=\frac{16}{3}-\frac{4}{3}H^2+\frac{8}{45}H^4-\frac{4}{315}H^6+\frac{8}{14175}H^8-\frac{8}{467775}H^{10}+O(H^{12}),$$

Implicit block method for $k=3$:

The implicit block methods for $k=3$, is given as in equations [2] and [3] in the previous section. In general form the equations can be written as below:

$$y_{n+1}=2y_n-y_{n-1}+h^2(P_3f_{n+1}+P_2f_n+P_1f_{n-1}+P_0f_{n-2}), \qquad [14]$$

$$y_{n+2}=2y_n-y_{n-2}+h^2(Q_4f_{n+2}+Q_3f_{n+1}+Q_2f_n+Q_1f_{n-1}). \qquad [15]$$

Substituting $y_n=\cos(\omega t_n)$ and $y_n=\sin(\omega t_n)$ respectively into [14] qand [15], resulting in,

$$2\cos(H)=2-H^2\big(P_3\cos(H)+P_2+P_1\cos(H)+P_0\cos(2H)\big), \qquad [16]$$

$$2\cos(2H)=2-H^2\big(Q_4\cos(2H)+Q_3\cos(H)+Q_2+Q_1\cos(H)\big), \qquad [17]$$

$$H^2\big(-P_3\sin(H)+P_1\sin(H)+P_0\sin(2H)\big)=0, \qquad [18]$$

$$H^2\left(-Q_4\sin(2H) - Q_3\sin(H) + Q_1\sin(H)\right) = 0. \qquad [19]$$

Solving equations [16 - [19] by letting P_2, P_3, Q_1 and Q_2 as the original values ($P_2 = \frac{5}{6}, P_3 = \frac{1}{12}, Q_1 = \frac{4}{3}, Q_2 = \frac{4}{3}$) and rewriting in Taylor series expansion, we have

$$P_0 = \frac{1}{240}H^4 - \frac{11}{60480}H^6 + \frac{13}{3628800}H^8 - \frac{1}{23950080}H^{10} + O(H^{12}),$$

$$P_1 = \frac{1}{12} - \frac{1}{120}H^4 + \frac{137}{30240}H^6 - \frac{139}{259200}H^8 + \frac{13}{427680}H^{10} + O(H^{12}),$$

$$Q_3 = \frac{4}{3} + \frac{2}{15}H^4 - \frac{16}{189}H^6 + \frac{1763}{113400}H^8 - \frac{1097}{748440}H^{10} + O(H^{12}),$$

$$Q_4 = -\frac{1}{15}H^4 + \frac{17}{1890}H^6 - \frac{113}{226800}H^8 + \frac{7}{427680}H^{10} + O(H^{12}).$$

Explicit block methods for $k = 5$:

The first and second point of the explicit block methods in Mansor et. al (2017) are given as follows:

$$y_{n+1} = 2y_n - y_{n-1} + h^2\left(\frac{299}{240}f_n - \frac{11}{15}f_{n-1} + \frac{97}{120}f_{n-2} - \frac{2}{5}f_{n-3} + \frac{19}{240}f_{n-4}\right). \qquad [20]$$

$$y_{n+2} = 2y_n - y_{n-2} + h^2\left(\frac{121}{15}f_n - \frac{184}{15}f_{n-1} + \frac{206}{15}f_{n-2} - \frac{104}{15}f_{n-3} + \frac{7}{5}f_{n-4}\right) \qquad [21]$$

In general form it can be written as

$$y_{n+1} = 2y_n - y_{n-1} + h^2(r_4 f_n + r_3 f_{n-1} + r_2 f_{n-2} + r_1 f_{n-3} + r_0 f_{n-4}), \qquad [22]$$

$$y_{n+2} = 2y_n - y_{n-2} + h^2(s_4 f_n + s_3 f_{n-1} + s_2 f_{n-2} + s_1 f_{n-3} + s_0 f_{n-4}). \qquad [23]$$

Substitute $y_n = \cos(\omega t_n)$ and $y_n = \sin(\omega t_n)$ respectively into [22] and [23], gives

$$2\cos(H) = 2 - H^2\left(r_4 + r_3\cos(H) + r_2\cos(2H) + r_1\cos(3H) + r_0\cos(4H)\right), \qquad [24]$$

$$2\cos(2H) = 2 - H^2\left(s_4 + s_3\cos(H) + s_2\cos(2H) + s_1\cos(3H) + s_0\cos(4H)\right), \qquad [25]$$

$$H^2\left(r_3\sin(H) + r_2\sin(2H) + r_1\sin(3H) + r_0\sin(4H)\right) = 0, \qquad [26]$$

$$H^2\left(s_3\sin(H) + s_2\sin(2H) + s_1\sin(3H) + s_0\sin(4H)\right) = 0. \qquad [27]$$

Then, solving equations [24] - [27] by letting r_0, r_1, r_4, s_0, s_1 and s_2 as the original values, we obtain

$$r_2=\frac{97}{120}-\frac{3}{40}H^4+\frac{787}{60480}H^6-\frac{1789}{1814400}H^8+\frac{10649}{239500800}H^{10}+O(H^{12}),$$

$$r_3=-\frac{11}{15}+\frac{3}{40}H^4-\frac{283}{30240}H^6+\frac{131}{259200}H^8-\frac{197}{11975040}H^{10}+O(H^{12}),$$

$$s_3=-\frac{184}{15}-\frac{22}{15}H^4+\frac{74}{45}H^6-\frac{221}{600}H^8+\frac{1181}{28350}H^{10}-\frac{40661}{13608000}H^{12}+O(H^{14}),$$

$$s_4=\frac{121}{15}+\frac{22}{15}H^4-\frac{229}{945}H^6+\frac{2041}{113400}H^8-\frac{859}{1069200}H^{10}+O(H^{12}).$$

Implicit block methods for $k = 5$:

The general form of the first and second point implicit block methods given in equations [4] and [5] in the previous section can be written as

$$y_{n+1}=2y_n-y_{n-1}+h^2(R_5f_{n+1}+R_4f_n+R_3f_{n-1}+R_2f_{n-2}+R_1f_{n-3}+R_0f_{n-4}),\quad [28]$$

$$y_{n+2}=2y_n-y_{n-2}+h^2(S_6f_{n+2}+S_5f_{n+1}+S_4f_n+S_3f_{n-1}+S_2f_{n-2}+S_1f_{n-3}).\quad [29]$$

Substituting $y_n=\cos(\omega t_n)$ and $y_n=\sin(\omega t_n)$ respectively into [28] and [29], resulting in

$$2\cos(H)=2-H^2\big(R_5\cos(H)+R_4+R_3\cos(H)+R_2\cos(2H)+R_1\cos(3H)+R_0\cos(4H)\big),\quad [30]$$

$$2\cos(2H)=2-H^2\big(S_6\cos(2H)+S_5\cos(H)+S_4+S_3\cos(H)+S_2\cos(2H)+S_1\cos(3H)\big),\quad [31]$$

$$H^2\big(-R_5\sin(H)+R_3\sin(H)+R_2\sin(2H)+R_1\sin(3H)+R_0\sin(4H)\big)=0,\quad [32]$$

$$H^2\big(-S_6\sin(2H)-S_5\sin(H)+S_3\sin(H)+S_2\sin(2H)+S_1\sin(3H)\big)=0.\quad [33]$$

Solving equations [30] - [33] by letting $R_0, R_2, R_3, R_4, S_1, S_3, S_4$ and S_5 as the original values we obtained values of R_1, R_5, S_2 and S_6 as follows:

$$R_1=-\frac{1}{40}+\frac{473}{241920}H^6-\frac{24223}{7257600}H^8+\frac{370289}{68428800}H^{10}+O(H^{12}),$$

$$R_5=\frac{3}{40}+\frac{137}{80640}H^6+\frac{7823}{2419200}H^8+\frac{57149}{10644480}H^{10}+O(H^{12}),$$

$$S_2=\frac{1}{15}+\frac{1}{945}H^6+\frac{221}{113400}H^8+\frac{23971}{7484400}H^{10}+\frac{212498413}{40864824000}H^{12}+O(H^{14}),$$

$$S_6 = \frac{1}{15} + \frac{1}{945}H^6 + \frac{221}{113400}H^8 + \frac{23971}{7484400}H^{10} + \frac{212498413}{40864824000}H^{12} + O(H^{14}).$$

RESULTS AND DISCUSSION

The proposed methods were implemented using predictor-corrector technique with only one iteration. The 2-point trigonometrically-fitted block explicit method for $k = 3$ is taken as the predictor equation and the 2-point trigonomerically-fitted block implicit method for $k = 3$ as the corrector equation this pair is denoted as two-point trigonometrically-fitted method of order 4 (TF2PBM4). The same goes for $k = 5$, where the 2-point trigonometrically-fitted block explicit method acts as the predictor and the 2-point trigonometrically-fitted block implicit method as the corrector, this pair is denoted as 2-point trigonometrically-fitted method of order 6 (TF2PBM6). We solved five tested problems that were obtained from the literature. Total time taken and maximum error would be shown in the form of efficiency curves.

Problem 1 [Rabiei et al., 2012]

$$y''(x) = -y(x), \qquad y(0) = 0, \qquad y'(0) = 1,$$

and the fitted frequency, ω=1. Exact solution is $y(x) = \sin(x)$.

Problem 2 [Jikantoro et al., 2015a]

$$y''(x) = -100y(x), \qquad y(0) = 1, \qquad y'(0) = -2,$$

and the fitted frequency, ω=10. Exact solution is $y(x) = -\frac{1}{5}\sin(10x) + \cos(10x)$.

Problem 3 [Senu et al., 2015]

$$y''(x) = -\omega^2 y(x) + (\omega^2 - 1)\sin(x), \qquad y(0) = 1, \qquad y'(0) = (\omega + 1),$$

and the fitted frequency, ω=10. Exact solution is $y(x) = \cos(\omega x) + \sin(\omega x) + \sin(x)$.

Problem 4 [Simos, 2003]

$$y_1''(x) = -y_1(x) + \epsilon\cos(\psi x), \qquad y_1(0) = 1, \qquad y_1'(0) = 0,$$

$$y_2''(x) = -y_2(x) + \epsilon\sin(\psi x), \qquad y_2(0) = 0, \qquad y_2'(0) = 1,$$

where $\epsilon = 0.001, \psi = 0.01$ and the estimated frequency, ω=1. Exact solutions are:

$$y_1(x) = \frac{1 - \epsilon - \psi^2}{1 - \psi^2}\cos(x) + \frac{\epsilon}{1 - \psi^2}\cos(\psi x),$$

$$y_2(x) = \frac{1 - \epsilon\psi - \psi^2}{1 - \psi^2}\sin(x) + \frac{\epsilon}{1 - \psi^2}\sin(\psi x).$$

Problem 5 [Jikantoro et al., 2015b]

$$y_1''(x) = -y_1(x) + 0.001\cos(x), \qquad y_1(0) = 1, \qquad y_1'(0) = 0,$$

$$y_2''(x) = -y_2(x) + 0.001\sin(x), \qquad y_2(0) = 0, \qquad y_2'(0) = 0.9995$$

and the fitted frequency, ω=1. Exact solutions are $y_1(x) = \cos(x) + 0.0005x\sin(x)$ and $y_2(x) = \sin(x) - 0.0005x\cos(x)$.

The notations used are as follows:

ω	Frequency of the problem
h	Step size
TIME(s)	Time taken to compute the method in second
MAXERR	Maximum error
TF2PBM4	The fourth order trigonometrically fitted 2-point block multistep method derived in this paper
TF2PBM6	The sixth order trigonometrically fitted 2-point block multistep method derived in this paper
ETSHMs	The fourth order explicit two-step hybrid method by Franco (2006).
IRKNM	The fourth order improved Runge-Kutta-Nystrom method with three stages by Rabiei et al. (2012).
PFRKN	The fourth order phase fitted Runge-Kutta-Nystrom method by Papadopoulos et al. (2008).
MSHMs	The four-step multistep hybrid method by Li and Wang (2016).
ETSHM6	The sixth order explicit two-step hybrid method by Franco (2006).
PFHM6	The sixth order phase fitted hybrid method by Senu et al. (2015).
NTM6	The sixth order explicit Numerov-type method by Tsitouras (2003).

The above methods are chosen as comparison because those are the methods usually used by most reserchers who are working on numerically solving oscillatory problems. The maximum error is defined by

$$MAXERR = \max|y(x_n) - y_n|,$$

where $y(x_n)$ is the exact solution and y_n is the approximate solution.

Methods of the same orders or steps were compared for the integration intervals of [1, 1000]. for methods of order four the efficiency curves for problems 1-5 are given in Figures 3 -7 and for methods of order six, the efficiency curves are given in Figures 8-12.

Five tested problems have been solved using the trigonometrically fitted block methods TF2PBM4 and TF2PBM6 of order four and five respectively. The maximum error of the

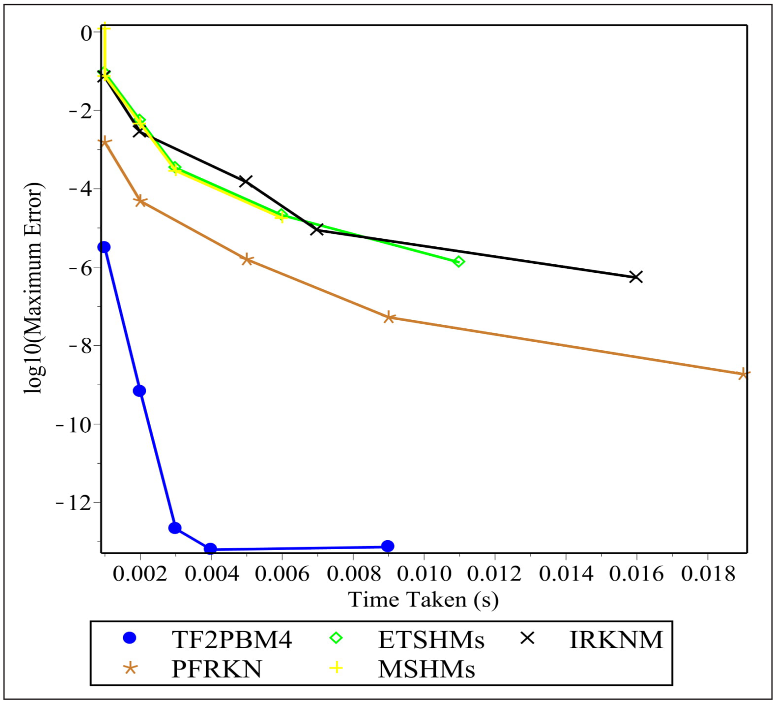


Figure 3. Efficiency curves of TF2PBM4 for Problem 1 with $xn = 1,000$ and $h = 0.5 \,/\, 2\text{^}i\,, i = 0, \ldots, 4$

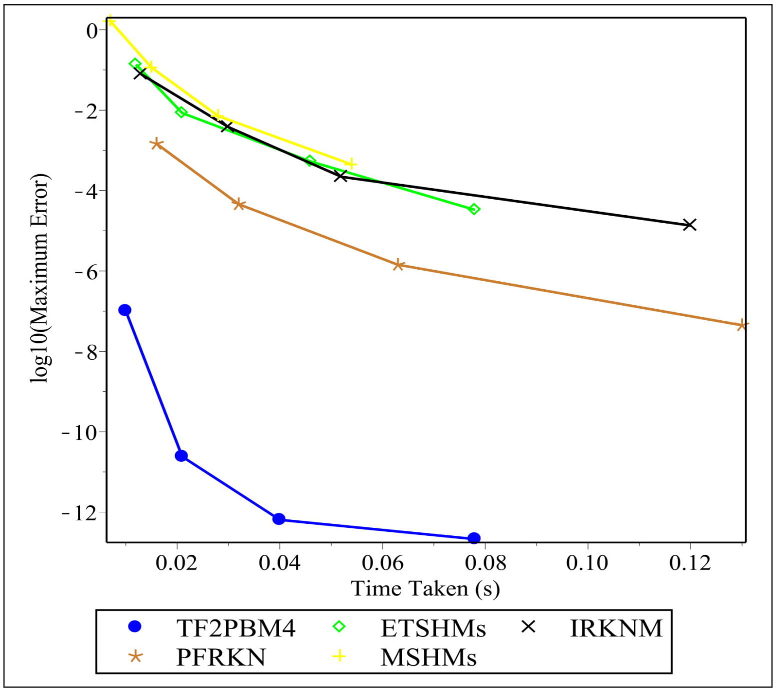


Figure 4. Efficiency curves of TF2PBM4 for Problem 2 with $xn = 1,000$ and $h = 0.125/2^{i}\,, i = 2, \ldots, 5$

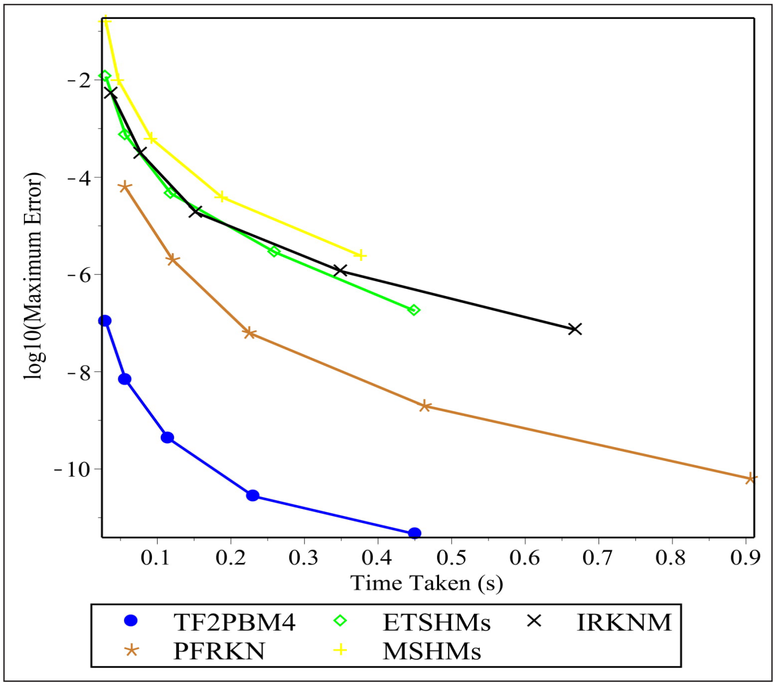


Figure 5. Efficiency curves of TF2PBM4 for Problem 3 with $xn = 1,000$ and $h = 0.125/2^i, i = 3, \ldots, 7$

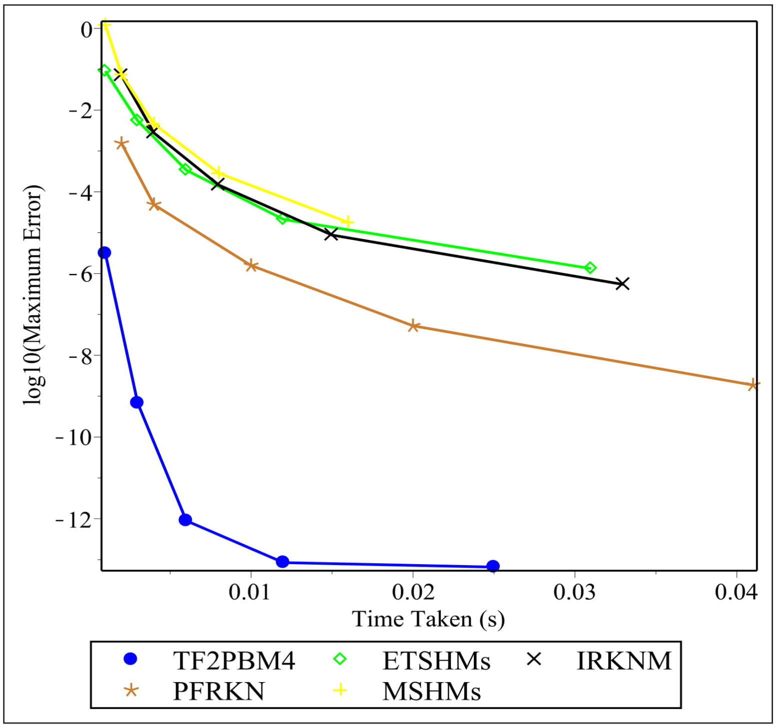


Figure 6. Efficiency curves of TF2PBM4 for Problem 4 with $xn = 1,000$ and $h = 0.5/2^i, i = 0, \ldots, 4$

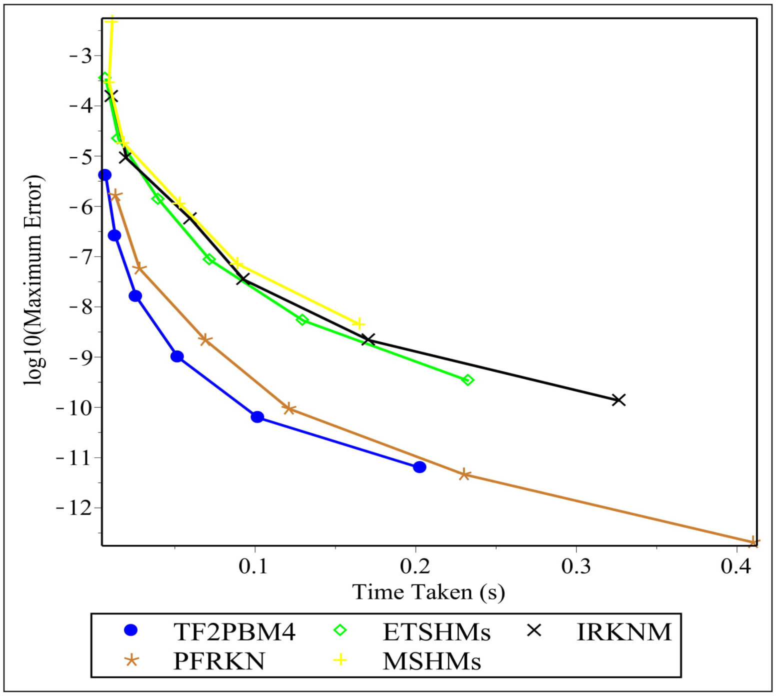


Figure 7. Efficiency curves of TF2PBM4 for Problem 5 with $xn = 1,000$ and $h = 0.125/2^i, i = 0, \ldots, 5$

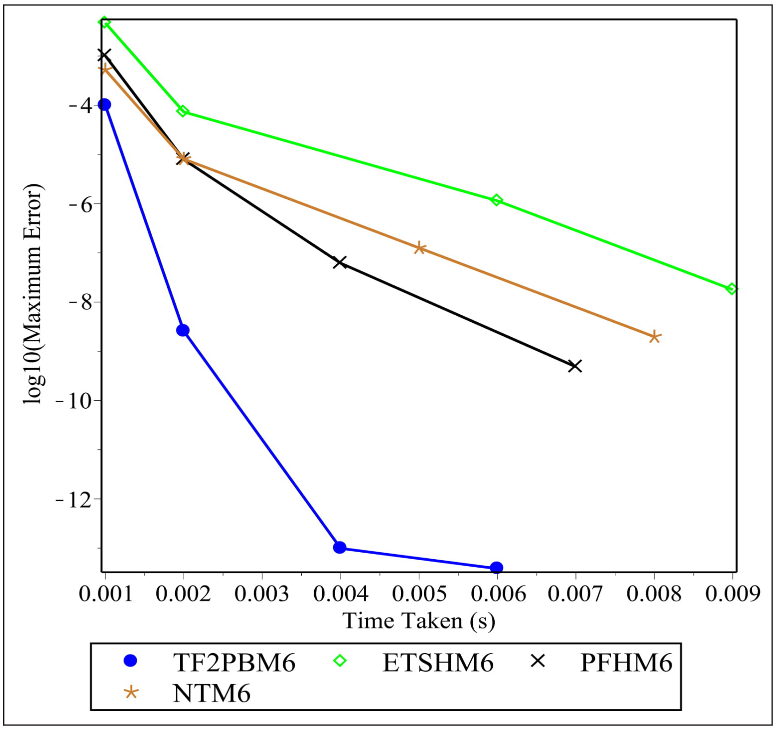


Figure 8. Efficiency curves of TF2PBM6 for Problem 1 with $xn = 1,000$ and $h = 0.5/2^i, i = 0, \ldots, 3$

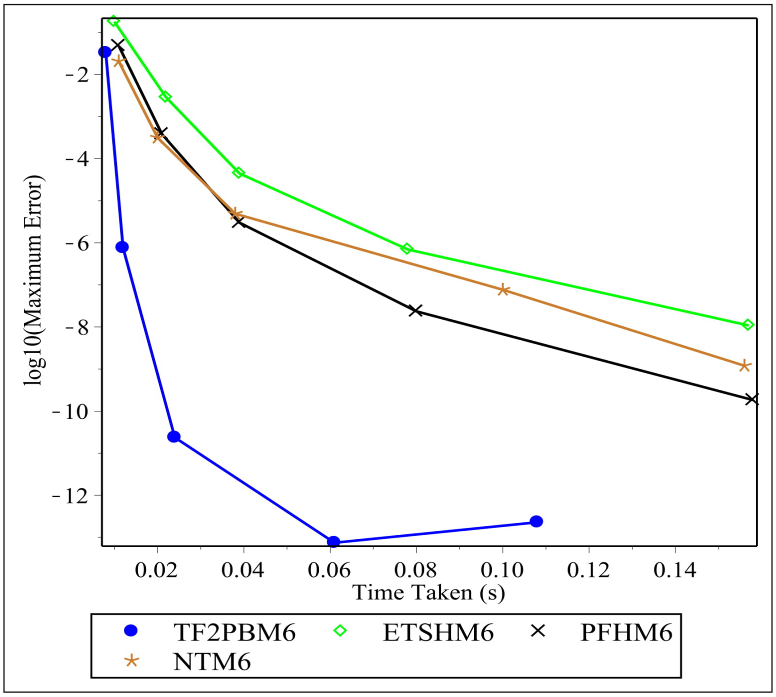


Figure 9. Efficiency curves TF2PBM6 for Problem 2 with $xn = 1,000$ and $h = 0.125/2^i, i = 1, \ldots, 5$

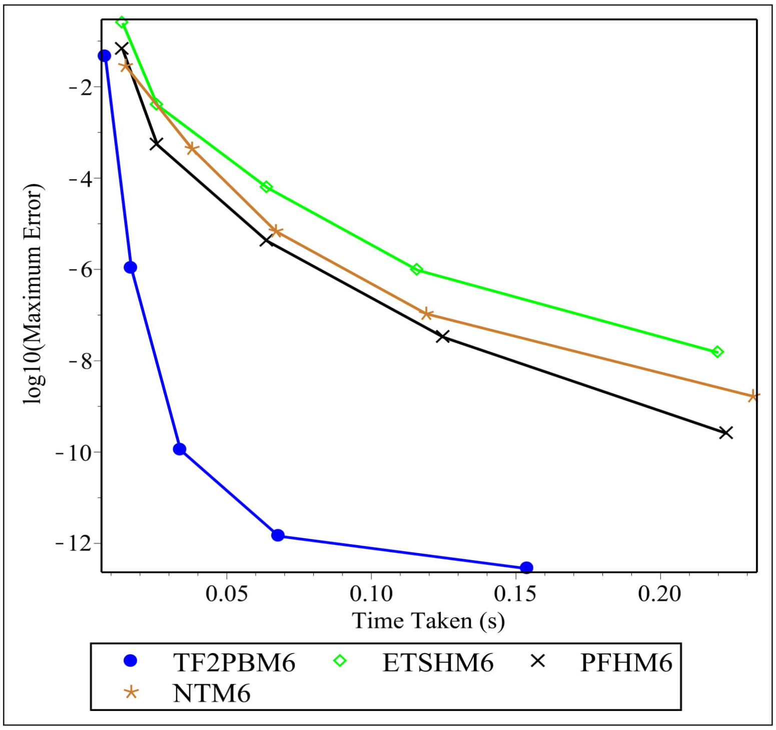


Figure 10. Efficiency curves TF2PBM6 for Problem 3 with $xn = 1,000$ and $h = 0.125/2^i, i = 3, \ldots, 7$

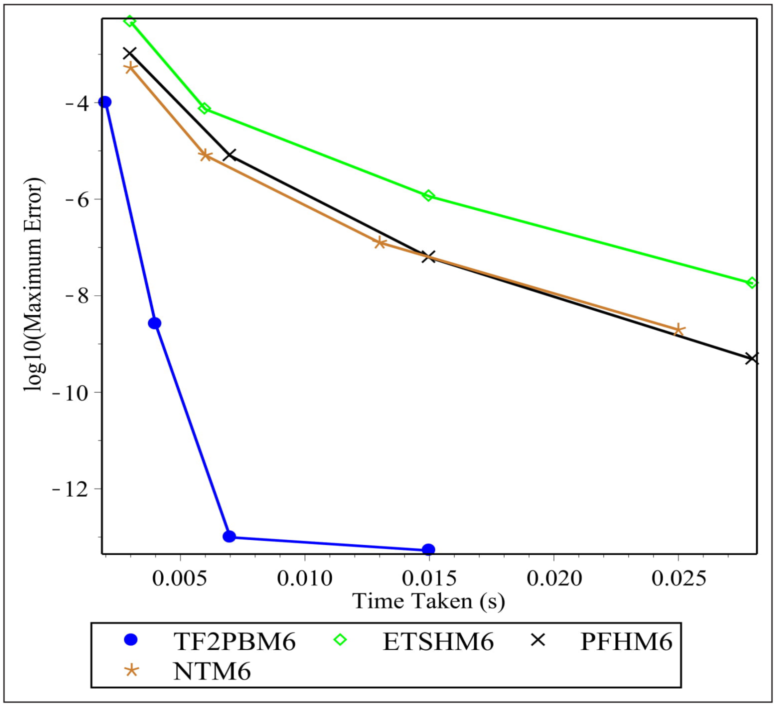


Figure 11. Efficiency curves TF2PBM6 for Problem 4 with $xn = 1,000$ and $h = 0.5/2^i, i = 0, \dots, 3$

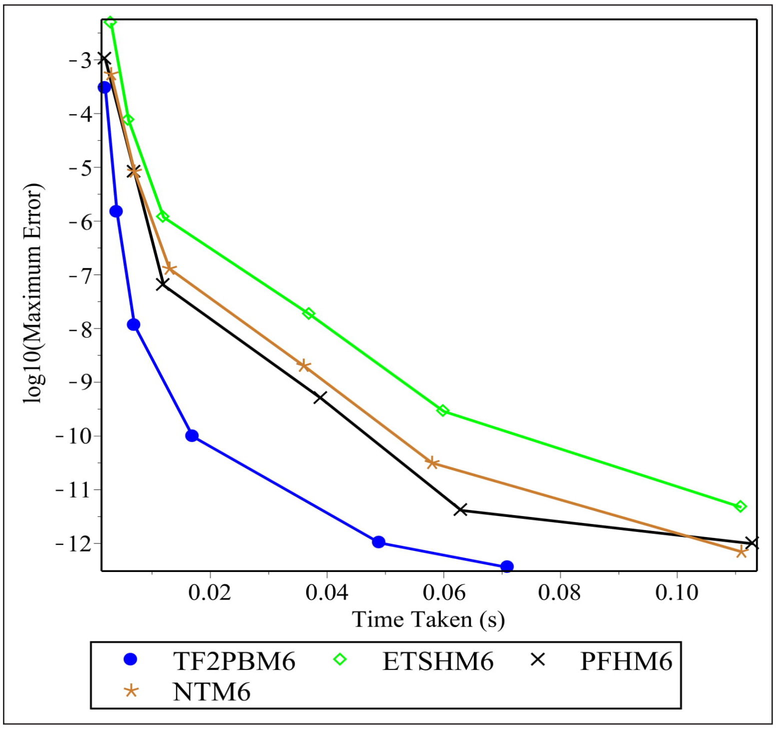


Figure 12. Efficiency curves TF2PBM6 for Problem 5 with $xn = 1,000$ and $h = 0.125/2^i, i = 0, \dots, 5$

new methods are plotted against execution time and they are compared with the existing methods based on the order of the methods. Based on the efficiency curves presented, for the fourth order methods, even though the execution time taken by MSHMs is shorter than TF2PBM4, TF2PBM4 gives more accurate results compared tot MSHMs. It is observed also that the execution time taken by TF2PBM6 is the shortest compared to the other methods. TF2PBM4 and TF2PBM6 have the smallest maximum error indicating that the new methods are more accurate compared to the existing methods. However, it can be observed that for certain problems, when the value of h is too small, the maximum error became larger as shown in Figure 9. This is because the value of $H = \omega h$ will approach zero when the value of h is too small and hence the coefficients of the trigonometrically fitted methods will approach the original methods.

CONCLUSION

In this paper, the 2-point block multistep methods that have been derived in Mansor et al. (2017) is shown to be absolutely stable. The methods are trigonometrically-fitted so that they are suitable for solving oscillatory problems. Codes based on the methods are developed using C Programming Language and are used to solve all the problems. The numerical results are compared with the existing methods in the scientific literature to present the performance of the proposed methods in the form efficiency curves. In conclusions, the new methods are superior than the existing methods in terms of accuracy and execution time. Hence trigonometric-fitting approach, enhanced the performance of the methods when used for solving oscillatory problems.

ACKNOWLEDGEMENT

We gratefully acknowledged Universiti Putra Malaysia for the financial assistance received through Putra Research Grant vote number 9543500**.**

REFERENCES

Ahmad, S. Z., Ismail, F., & Senu, N. (2016). A four-stage fifth-order trigonometrically fitted semi-implicit hybrid method for solving second-order delay differential equations. *Mathematical Problems in Engineering, 2016*, 1-7.

Akinfenwa, O., Akinnukawe, B., & Mudasiru, S. (2015). A family of continuous third derivative block block methods for solving stiff systems of first order ordinary differential equations. *Journal of the Nigerian Mathematical Society,* 34(2), 160-168.

Fang, Y., & Wu, X. (2008). A trigonometrically fitted explicit numerov-type method for second-order initial value problems with oscillating solutions. *Applied Numerical Mathematics, 58*, 341-351.

Fatunla, S. O. (1995). A Class of Block Methods for Second Order IVPs. *International Journal of Computer Mathematics, 55*(1-2), 119-133.

Franco, J. M. (2006). A class of explicit two-step hybrid methods for second-order IVPs. *Journal of Computational and Applied Mathematics, 187*, 41-57.

Jikantoro, Y. D., Ismail, F., & Senu, N. (2015a). Zero-dissipative semi-implicit hybrid method for solving oscillatory or periodic problems. *Applied Mathematics and Computation, 252*, 388-396.

Jikantoro, Y. D., Ismail, F., & Senu, N., (2015b). Zero-dissipative trigonometrically fitted hybrid method for numerical solution of oscillatory problems. *Sains Malaysiana, 44*(3), 473-482.

Li, J., & Wang, X. (2016). Multi-step hybrid methods for special second-order differential equations $y''(t) = f(t, y(t))$. *Numerical Algorithms, 73*(3), 711-733.

Li, J., Lu, M., & Qi, X. (2017). Trigonometrically fitted multistep hybrid methods for oscillatory special second order initial value problems. *International Journal of Computer Mathematics*, *95*(5), 979-997.

Mansor, A. F., Ismail, F., & Senu, N. (2017). Explicit and implicit block multistep methods for solving special second order ordinary differential equations. *International Journal of Applied and Engineering Research, 12*(24), 14176-14189.

Papadopoulos, D. F., Anastassi, Z. A., & Simos, T. E. (2008). A phase-fitted runge-kutta-nyström method for the numerical solution of initial value problems with oscillating solutions. *Computer Physics Communications, 180*, 1839-1846.

Rabiei, F., Ismail, F., Senu, N., & Abasi, N. (2012). Construction of improved runge-kutta-nystrom method for solving second-order ordinary differential equations. *World Applied Science Journal, 20*, 1685-1695.

Ramos, H., Kalogiratou, Z., Monovasilis, T., & Simos, T. E. (2016). An optimized two-step hybrid block method for solving general second order initial-value problems. *Numerical Algorithms, 72*(4), 1089-1102.

Senu, N., Ismail, F., Ahmad, S. Z., & Suleiman, M. (2015). Optimize hybrid methods for solving oscillatory second order initial value problems. *Discrete Dynamics in Nature and Society, 2015*, 1-11.

Simos, T. E. (2003). Exponentially-fitted and trigonometrically-fitted symmetric linear multistep methods for the numerical integration of orbital problems. *Physics Letters A, 315*, 437-446.

Tsitouras, C. H. (2003). Explicit numerov type methods with reduced number of stages. *Computer and Mathematics with Applications, 45*, 37-42.

Pertanika J. Sci. & Technol. 27 (3): 2399 - 2408 (2019)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Effect of Parboiling on *in vitro* Physiological Antioxidant Capacity of Brown Rice

Elaine Lee, Phene Neoh Pei Nee, Lee Huei Hong, Wong Sie Chuong, Wong Tze Jin, Koo Lee Feng and Yiu Pang Hung*

Faculty of Agriculture and Food Sciences, Universiti Putra Malaysia, 97000 UPM, Bintulu, Sarawak, Malaysia

ABSTRACT

Parboiling process has been widely implemented in brown rice processing, but its effect on *in vitro* physiological antioxidant capacity of brown rice was not known. In this study, an *in vitro* method simulating the human physiological conditions was used to investigate the effect of parboiling on antioxidant capacity of brown rice in three Bario rice varieties. In this method, bacterial inocula were prepared from rat cecal contents. Results showed that parboiling process gave significant impacts on *in vitro* physiological antioxidant capacity of brown rice. The process improved total phenolic content at small intestine (Adan Halus), DPPH scavenging activity at both small and large intestines (Adan Halus and Bario Merah) and ferrous ion-chelating activity at large intestine (Bario Hitam). However, changes in antioxidant capacity were variety dependent, possibly due to different bran pigmentation. These suggested that parboiling process could improve physiological antioxidant capacity with *in vitro* simulation at small and large intestines by selecting a suitable rice variety and parboiled brown rice could offer good antioxidant properties to maintain physiological health.

Keywords: Bario rice, brown rice, DPPH scavenging activity, ferrous chelating activity, *in vitro* physiological antioxidant capacity, parboiled rice, total phenolic content

ARTICLE INFO

Article history:
Received: 27 August 2018
Accepted: 25 June 2019
Published: 21 October 2019

E-mail addresses:
ylang_329@hotmail.com (Elaine Lee)
pheneneoh@yahoo.com (Phene Neoh Pei Nee)
hueihong24@gmail.com (Lee Huei Hong)
wongsie@upm.edu.my (Wong Sie Chuong)
w.tzejin@upm.edu.my (Wong Tze Jin)
leefeng@upm.edu.my (Koo Lee Feng)
yiuph@upm.edu.my (Yiu Pang Hung)
*Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

INTRODUCTION

Whole grain rice or brown rice is a rich source of antioxidants for human diets. Significant scientific evidence indicates that consumption of whole grain products may reduce the risk for various types of chronic diseases (Liu et al., 2000). Brown rice is rich in phytochemicals such as derivatives of benzoic and cinnamic acids, anthocyanidins, flavonols and phenolic compounds. Some of these phytochemicals, such as ferulic acid, and diferulates are predominantly found in grains but are not present in significant quantities in fruits and vegetables (Bunzel et al., 2001). Due to the escalating health consciousness among the consumers, brown rice is gaining interest for its excellent health beneficial properties.

Nevertheless, the slow water absorption and rapid rancidity of brown rice lead to low acceptance among rice consumers. Parboiling process can strengthen the kernel integrity, prolong the shelf life of rice, shorten the cooking period, and prevent the loss of nutrients and the proliferation of fungus in rice (Bhattacharya, 2004). Besides, parboiling process also increases the nutritional values of the rice by induce nutrients from the bran into the kernel.

Several studies have reported on higher level of antioxidants in parboiled rice than non-parboiled rice (Bhattacharya, 2011; Byungrok et al., 2012). Antioxidant properties are commonly evaluated on solvent extracts of rice samples, which may not reflect the *in vivo* availability of the antioxidants (Serrano et al., 2007). Colonic microbial fermentation may also releases the additional antioxidants which are not possible to be evaluated using solvent extraction method alone (Serrano et al., 2007). Thus, it is important to evaluate the antioxidant release under physiological conditions to stimulate the human digestive system via gastric and intestinal digestion (Etcheverry et al., 2012).

Parboiling process could amend the drawbacks in brown rice and provide better nutritional quality. Brown rice could have different antioxidant capacity at small and large intestines after being parboiled. In this study, parboiling process was first applied on Bario brown rice to compare the *in vitro* physiological antioxidant capacity of non-parboiled and parboiled Bario brown rice at physiological conditions.

MATERIALS AND METHODS

Chemicals

The enzymes used for the digestive enzymatic treatment were pepsin (Acros), pancreatin (Sigma-Aldrich, from porcine pancreas), lipase (Sigma-Aldrich, from porcine pancreas), bile extract (Sigma-Aldrich, bile extract from porcine), α- amylase (Sigma-Aldrich, from *Aspergillus oryzae*), and amyloglucosidase (Fluka, from *Aspergillus niger*). Reagents used for preparing the buffer for digestive treatment were: hydrochloric acid (37%, Merck), potassium chloride (Merck), dipotassium phosphate (Scharlau), potassium phosphate monobasic (Riedel-de Haen), Tris (Vivantis), and maleic acid.

Chemicals for the colonic fermentation medium were ammonium bicarbonate (Merck), sodium bicarbonate (Fluka), disodium phosphate, anhydrous (Merck), monopotassium phosphate, anhydrous (Merck), magnesium sulfate heptahydrate, calcium chloride dehydrate (Merck), manganese (II) chloride tetrahydrate (Merck), cobalt (II) chloride hexahydrate (Merck), iron (III) chloride hexahydrate (Merck), resazurin sodium salt powder (Sigma-Aldrich), dithiothreitol (Merck), sodium sulfide monohydrate (Fisher), lactulose (Sigma-Aldrich), and trypticase peptone (pancreatic digest of casein).

Rice Samples

Three popular rice cultivars (*Oryza sativa* L.) of Sarawak, Malaysia namely Bario rice, Adan Halus, Bario Merah, and Bario Hitam were selected. The paddy (3 – 5 kg) of Bario rice varieties were collected from Bario Highland, Miri, Malaysia, de-husked and stored at -20°C until use (Lee et al., 2015).

Parboiling Treatment

Parboiling treatment at laboratory-scale was conducted according to modified parboiling method from Patindol and colleagues (2008). Full grains of brown rice (300 mg) were soaked in 120 mL water at 60°C for 4 hours. The soaked rice grains were then immersed in water bath at the same temperature for another one hour after the excess water was drained off. Later, the rice grains were autoclaved at 121 °C for 20 min at 15 psi. Finally, the rice grains were dried at 60°C for 45 minutes in an oven. The parboiling treatment was conducted with factorial arrangement in completely randomized design. Brown rice untreated with parboiling process was assigned as a control for comparison.

In vitro simulation of physiological conditions

In vitro physiological digestion and colonic fermentation were conducted according to Serrano et al. (2007).

Digestive Enzymatic Treatment. Parboiled and non-parboiled rice of three rice varieties were first incubated with digestive enzymes in three replications. Briefly, 300 mg of rice grains were crunched and incubated with pepsin (0.2mL of a 300mg / mL solution in HCl-KCl 0.2 M buffer, *pH 1.5, 40* °C, 1 h), pancreatin (1mL of a 5mg /mL solution in phosphate buffer 0.1 M, pH 7.5, 37 °C, 6 h), lipase (2mL of a 7 mg/mL solution in phosphate buffer 0.1 M, pH 7.5, 37 °C, 6 h), bile extract porcine (2mL of a 17.5 mg/mL solution in phosphate buffer 0.1 M, pH 7.5, 37 °C, 6 h) and α-amylase (1mL of a 120mg/mL solution in tris-maleate buffer 0.1 M, pH 6.9, 37 °C, 16 h). After the incubation period, the samples were centrifuged for 15 minutes at 25 °C (3000 g). Upon removal of the supernatants, pellet residues for each sample were washed twice with 5 mL of distilled water and all the supernatants were combined. The collected supernatants were incubated with 100 µL of amyloglucosidase for 45 minutes at 60 °C. The supernatants were then stored at -20 °C for antioxidant capacity analysis. Blanks containing no substrate and lactulose were included in the experiment for negative and positive control respectively.

Colonic Fermentation. The pellet residues collected were stored in at -20 °C prior to *in vitro* colonic fermentation. Male rats (body weight of 100 ± 5 g) were killed in a carbon dioxide chamber. Fresh rat cecum was collected through the abdominal midline incisions of rat's body (AUP-R049/2013). The rat cecal contents was scraped, weighed and added to a flask containing sterile anaerobic medium to give a 100 g/L inoculum. The anaerobic medium was prepared according to Goering and Van (1970) which contained trypticase, micromineral and macromineral solutions and also the resazurin as anaerobic redox indicator. The indigestible fractions (pellet residues) collected from the digestive enzymatic treatment were mixed with fermentation medium (8 mL, 4 °C, 16 h) followed by addition of 2 mL inoculums. The headspace of test tubes were rinsed with carbon dioxide before sealed with parafilm. The colonic fermentation was conducted at 37°C for 24 hours. Blanks containing no substrate and lactulose were included in the experiment to function as zero and completely fermentable substrate, respectively. After the incubation period, pH values of the samples were measured and neutralized with 1 M NaOH. The test tubes with fermentation mixtures were centrifuged (2500g, 10 min, 25°C). The supernatants were collected and stored at -20°C for antioxidant capacity analysis.

In vitro Physiological Antioxidant Capacity

Antioxidant capacity was measured on the supernatant collected from digestive enzymatic treatment and colonic fermentation. The supernatant collected contains the physiological available fractions of the chemical compounds, which contributes to antioxidant capacity physiologically. Three antioxidant capacity, namely phenolic content, DPPH scavenging activity and ferrous ion chelating activity were measured.

Phenolic Content. Phenolic content was determined from supernatants collected according to Butsat and Siriamornpun (2010). An aliquot of 80 µL supernatant was mixed with 400 µL of freshly prepared Folin-Ciocalteu reagent (0.2 N). Then, 320 µL of sodium carbonate (Na_2CO_3) and 600 µL of ultrapure water were added. The mixtures were then topped up to 1 mL and left to stand for 2 hours at room temperature. Absorbance at 760 nm was measured by using a UV-vis spectrophotometer. Phenolic content was expressed as mg gallic acid equivalent per gram sample.

DPPH Radical Scavenging Activity. DPPH radical scavenging activity of the supernatants was determined according to Bran-Williams et al. (1995). Supernatant (0.5 mL) was mixed with 2.5 mL of a 0.5 mM methanolic DPPH solution. The mixture was then shaken vigorously and incubated for 30 minutes in the dark at room temperature. The absorbance at 517 nm was measured by using a UV-vis spectrophometer. DPPH free radical-scavenging ability was calculated by using the formula: DPPH scavenging activity (%) = [Absorbance of control-Absorbance of sample / Absorbance of control] ×100.

Ferrous Ion Chelating Activity. Ferrous ion-chelating activity of the supernatants was measured according to Zhao et al. (2008). Supernatant (50 µL) was mixed with 50 µL of ferrous chloride solution. Then, 1.6 mL of 80% methanol was added to the mixture and incubated for 5 minutes at room temperature. Ferrozine (100 µL) was then added and further incubated for 10 minutes at room temperature. The absorbance at 562 nm was measured by using a UV-vis spectrophometer. Ferrous ion-chelating ability was calculated by using the formula: (%) = [1-Absorbance of sample / Absorbance of control] ×100. Disodium EDTA had been used as a standard.

Data Analysis

Data was analyzed statistically using independent t-test to detect the differences of antioxidant capacity between parboiled and non-parboiled rice, and between small and large intestines ($p<0.05$). The statistical software used was Statistical Analysis System (SAS) version 9.3.

RESULTS AND DISCUSSION

Total Phenolic Content

Brown rice or non-parboiled rice of “Adan Halus” and “Bario Hitam” showed higher bioaccessibility of phenolic compounds at large intestine (Table 1). After parboiling treatment, these rice samples showed higher phenolic content at small intestine than the large intestine. The phenolic content of parboiled “Adan Halus” at small intestine increased up to 69.79 % (8.30 ±1.53 to 14.09 ± 0.87 µg gallic acid equivalents/g rice) compared to non-parboiled “Adan Halus”. The phenolic content of “Bario Hitam” at small intestine increased 22.05 % compared to non-parboiled rice. Parboiled “Bario Merah” showed 24.83 % loss in phenolic content compared to non-parboiled “Bario Merah”. The loss of phenolic content was also accompanied by the change in bran color after parboiling (Figure 1). “Bario Merah” could contained high level of readily soluble or unbound antioxidants which then contributed to the higher extractability of phenolic compounds at small intestine.

The higher concentration of phenolic content in the small intestine shown by “Adan Halus” and “Bario Hitam” indicated the release of bounded compounds from the originally complexes through the parboiling process. Higher concentration of phenolic content at

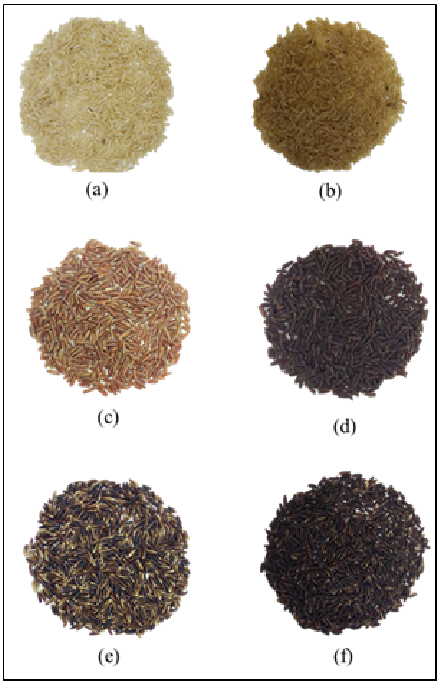


Figure 1. The diagram showed parboiled (right) and non-parboiled (left) rice samples for three rice varieties: "Adan Halus" (a&b), "Bario Merah" (c&d) and "Bario Hitam" (e&f).

small intestine was an important turn up as most of the polyphenols get absorbed in this site (Pandey & Rizvi 2009). Although the high concentration of antioxidants did not guarantee greater absorption of antioxidants at the small intestine, it was possible that the cellular uptake of metabolites is proportional to their unbound concentration. Hence, the possible uptake might enhance the function of antioxidants in human's body in preventing oxidative stress related diseases.

The effect of parboiling on overall phenolic content from small intestine until large intestine was not consistent among rice varieties. Overall phenolic content of "Adan Halus" and "Bario Hitam" increased after parboiling but not in "Bario Merah". The increment in overall phenolic content of parboiled rice was in congruence to previous study which reported on the release of phenolic compounds which were once bounded to the cell wall (Min 2006). However, different phenolics profiles were expected in "Adan Halus", "Bario Hitam" and "Bario Merah". Thermal energy in the presence of oxygen and moisture would accelerate the oxidative degradation of phenolic compounds and led to destruction of some phenolic compounds specifically present in Bario Merah. This could explained the lower overall phenolic content of "Bario Merah", which may due to the possible loss of proanthocyanidins with low degree polymerization (Min, 2006; Awika et al., 2003).

DPPH Scavenging Activity

DPPH scavenging activity of brown rice or non-parboiled rice of three rice varieties was higher in small intestine than large intestine (Table 2). After parboiling, DPPH scavenging activity was increased at both small and large intestine following the same trend of higher DPPH scavenging activity in small intestine. The increment of DPPH radical scavenging activity in all rice varieties was probably caused by detachment of the bounded antioxidants other than phenolic compounds from the cell wall after the thermal activity in parboiling process. These could be tocopherols, tocotrienols and γ-oryzanol (Iqbal et al., 2005). The increment of DPPH scavenging activity of brown rice after parboiled in

Table 1

Total phenolic content of parboiled and non-parboiled Bario rice varieties.

	Adan Halus		Bario Merah		Bario Hitam	
	Non-parboiled	Parboiled	Non-parboiled	Parboiled	Non-parboiled	Parboiled
Small intestine	8.30 ± 1.53y,b	14.09 ± 0.87x,a	18.02 ± 0.64x,a	13.54 ± 1.63x,b	13.00 ± 0.80y,a	16.74 ± 1.20x,a
Large intestine	5.06 ± 0.20x,a	4.56 ± 0.39y,b	4.96 ± 0.26y,a	4.31 ± 0.21y,b	5.28 ± 0.24x,a	4.47 ± 0.19y,a
Overall	13.35± 1.37	18.65 ± 0.54	22.97±0.40	17.85±1.50	18.28±0.85	21.22±1.20

Note.

1. Mean value ± standard deviation, n=3
2. Mean values followed by different letters of x and y within column show significant difference between small intestine and large intestine at $p<0.05$
3. Mean values followed by different letters of a and b within row show significant difference between non-parboiled and parboiled treatments of same variety at $p<0.05$

Table 2

DPPH scavenging activity of parboiled and non-parboiled Bario rice varieties.

	Adan Halus		Bario Merah		Bario Hitam	
	Non-parboiled	Parboiled	Non-parboiled	Parboiled	Non-parboiled	Parboiled
Small intestine	24.11 ± 0.72x,b	26.96 ± 1.76x,a	20.76 ± 3.58x,b	24.89 ± 2.34x,a	34.88 ± 2.68x,a	40.78 ± 1.80x,a
Large intestine	14.03 ± 2.15y,b	16.63 ± 1.45y,a	10.21 ± 0.62y,b	24.10 ± 1.84y,a	10.43 ± 2.68y,a	9.27 ± 0.43y,a
Overall	19.07±1.12	21.80±1.54	15.49±1.93	24.50±1.88	22.65±1.98	25.02±1.08

Note.

1. Mean value ± standard deviation, n=3
2. Mean values followed by different letters of x and y within column show significant difference between small intestine and large intestine at $p<0.05$
3. Mean values followed by different letters of a and b within row show significant difference between non-parboiled and parboiled treatments of same variety at $p<0.05$

Table 3

Ferrous chelating activity of parboiled and non-parboiled Bario rice varieties.

	Adan Halus		Bario Merah		Bario Hitam	
	Non-parboiled	Parboiled	Non-parboiled	Parboiled	Non-parboiled	Parboiled
Small intestine	19.79 ± 7.96y,a	N.D y,a	N.D. y,a	N.D. y,a	16.82 ± 4.13x,b	19.54 ± 5.51y,a
Large intestine	79.61 ± 38.94x,a	92.85 ± 12.39x,a	94.40 ± 3.74x,a	97.11 ± 3.40x,a	15.77 ± 11.39x,b	96.05 ± 6.85x,a
Overall	56.31±5.28	46.42±6.20	47.20±1.87	48.55±1.70	16.29±7.22	57.79±4.63

Note.

1. Mean value ± standard deviation, n=3
2. Mean values followed by different letters of x and y within column show significant difference between small intestine and large intestine at $p<0.05$
3. Mean values followed by different letters of a and b within row show significant difference between non-parboiled and parboiled treatments of same variety at $p<0.05$
4. N.D. indicates non detectable

small and large intestines possess benefits for optimal antioxidant absorption and also maintenance of colon health. The free antioxidant compounds might counteract the pro-oxidant or toxic species produced during colonic bacterial metabolism (Serrano et al., 2007).

In general, parboiling treatment gave similar increment trend of DPPH scavenging activity on all rice varieties but at different magnitude. DPPH radical-scavenging activity of "Adan Halus" (5.45 %) and "Bario Merah" (18.02 %) increased significantly after parboiling, except "Bario Hitam". The increases in DPPH scavenging activity could be explained by release of antioxidant compounds from parboiling treatment (Brighente et al., 2007). However, there's no changes in "Bario Hitam" possibly due to the loss of anthocyanin during parboiling with observed discoloration of the grains. The loss of DPPH scavenging activity due to loss of anthocyanin might have been compensated with the release of other antioxidant compounds after parboiling, and resulted in no significant difference between parboiled and non-parboiled rice for "Bario Hitam".

Ferrous Chelating Activity

Ferrous chelating activity of brown rice or non-parboiled rice of three rice varieties was higher in large intestine than small intestine (Table 3). The ferrous chelating activity of parboiled rice at the large intestine was 92.85 ± 12.39 %, 97.11 ± 3.40 % and 96.05 ± 6.85 % for "Adan

Halus", "Bario Merah" and "Bario Hitam", respectively. This was most probably caused by the low stability of chelating agents in the low pH conditions before enter small intestine (Moon & Shibamoto, 2009). Ferrous ion-chelating activity detected in the small intestine is therefore lesser than the activity in large intestine.

After parboiling treatment, ferrous chelating activity was increased at both small and large intestine only in "Bario Hitam". The chelating activity of parboiled "Bario Hitam" showed significantly higher values than the non-parboiled "Bario Hitam" in the large intestine with the increment of approximately 80.28%. Besides, "Bario Hitam" could contain bounded antioxidants with catechol and galloyl groups (Perron & Brumaghim, 2009). Parboiling process successfully released these antioxidants and contribute to higher ferrous chelating activity in parboiled "Bario Hitam".

CONCLUSIONS

Parboiling treatment significantly affected the *in vivo* antioxidant capacity of Bario rice varieties base on the current simulation study. The effect of parboiling was variety dependent due to the differences in antioxidants composition. Parboiled rice showed improved phenolic content at small intestine ("Adan Halus"), DPPH scavenging activity at both small and large intestines ("Adan Halus" and "Bario Merah") and ferrous ion-chelating activity at large intestine ("Bario Hitam"). The study indicated that antioxidant capacity from parboiling process is varied across rice varieties.

ACKNOWLEDGEMENT

The authors gratefully acknowledge the research fund given by Universiti Putra Malaysia, Malaysia (9589200) and generous assistance from the Sarawak Department of Agriculture.

REFERENCES

Awika, J. M., Dykes, L., Gu, L., Rooney, L. W., & Prior, R. L. (2003). Processing of sorghum and sorghum products alters procyanidin oligomer and polymer distribution and content. *Journal of Agricultural and Food Chemistry, 51*(18), 5516-5521.

Bhattacharya, K. R. (2004). Parboiling of rice. In E. T. Champagne (Ed.), *Rice: chemistry and technology* (pp. 329-404). Louisiana, United States: Department of Agriculture Research Service, Southern Regional Research Center.

Bhattacharya, K. R. (2011). *Rice quality: A guide to rice properties and analysis*. Cambridge, United Kingdom: Woodhead Publishing Ltd.

Bran-William, W., Cuvelier, M. E., & Berset, C. (1995). Use of a free radical method to evaluate antioxidant activity. *Food Science and Technology, 28*, 25-30.

Brighente, I. M. C., Dias, M., Verdi, L. G., & Pizzolatti, M. G. (2007). Antioxidant activity and total phenolic content of some Brazilian species. *Pharmaceutical Biology, 45*(2), 156-61.

Bunzel, M., Ralph, J., Marita, J. M., Hatfield, R. D., & Steinhar, H. (2001). Diferulates as structural components in soluble and insoluble cereal dietary fibre. *Journal of the Science of Food and Agriculture, 81*, 653-660.

Butsat, S., & Siriamornpun, S. (2010). Antioxidant capacities and phenolic compounds of the husk, bran and endosperm of Thai rice. *Food Chemistry, 119*, 606-613.

Byungrok, M., Gu, L. W., Anna, M. M., Christine, J. B., & Chen, M. H. (2012). Free and bound total phenolic concentrations, antioxidant capacities, and profiles of proanthocyanidins and anthocyanins in whole grain rice (*Oryza sativa* L.) of different bran colours. *Food Chemistry, 133*, 715-722.

Etcheverry, P., Grusak, M. A., & Fleige, L. E. (2012). Application of *in vitro* bioaccessibility and bioavailability methods for calcium, carotenoids, folate, iron, magnesium, polyphenols, zinc, and vitamins B6, B12, D, and E. *Front Physiology, 3*, 1-21.

Goering, H. K., & Van, S. (1970). *Forage fiber analyses (Apparatus, reagents, procedures and some applications): Agriculture handbook No.379 Agricultural Research Service.* Washington DC, United States: Department of Agriculture.

Iqbal, M. J., Yaegashi, S., Ahsan, R., Shopinski, K. L., & Lightfoot, D. A. (2005). Root response to *Fusarium solani* f. sp. *glycines*: temporal accumulation of transcripts in partially resistant and susceptible soybean. *Theoretical and Applied Genetics, 110*, 1429-1438.

Lee, H. H., Loh, S. P., Bong, C. F. J., Sarbini, S. R., & Yiu, P. H. (2015). Impact of phytic acid on nutrient bioaccessibility and antioxidant properties of dehusked rice. *Journal of Food Science and Technology, 52*(12), 7806-7816.

Liu, S., Manson, J. E., Stampfer, M. J., Hu, F. B., Giovannucci, E., Colditz, G. A., ... & Willett, W. C. (2000). A prospective study of whole grain intake and risk of type 2 diabetes mellitus in U.S. women. *American Journal of Public Health, 90*, 1409-1415.

Min, B. (2006). Effects of hydrothermal processes on antioxidants in brown , purple and red bran whole grain rice *(Oryza sativa* L.). *Food Chemistry, 159*, 106-150.

Moon, J. K., & Shibamoto, T. (2009). Antioxidant assays for plant and food components. *Journal of Agricultural and Food Chemistry, 57*(5), 1655-1666.

Pandey, K. B., & Rizvi, S. I. (2009). Plant polyphenols as dietary antioxidants in human health and disease. *Oxidative Medicine and Cellular Longevity, 2*(5), 270-278.

Patindol, J., Newton, J., & Wang, Y. J. (2008). Functional properties as affected by laboratory-scale parboiling of rough rice and brown rice. *Journal of Food Science, 73*, 370-377.

Perron, N. R., & Brumaghim, J. L. (2009). A review of the antioxidant mechanisms of polyphenol compounds related to iron binding. *Cell Biochemistry and Biophysics, 53*(2), 75-100.

Serrano, J., Goñi, I., & Saura-Calixto, F. (2007). Food antioxidant capacity determined by chemical methods may underestimate the physiological antioxidant capacity. *Food Research International, 40*(1), 15-21.

Zhao, H., Fan, W., Dong, J., Lu, J., Chen, J., & Shan, L. (2008). Evaluation of antioxidant activities and total phenolic contents of typical malting barley varieties. *Food Chemistry, 107*, 296-304.

Pertanika J. Sci. & Technol. 27 (4): 2409 - 2425 (2019)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Characteristics of Chitosan Nanoparticles extracted from Sea Cucumber (*Holothuria scabra*) as Source Materials for Glucosamine

Sumarto*, Bustari Hasan, Rahman Karnila and Mery Sukmiwati

Department of Fishery Products Processing, Faculty of Fisheries and Marine Science, Riau University, Pekanbaru, Riau, 28293, Indonesia

ABSTRACT

Sea cucumber has a thick layer of skin consisting of lime components. In fact, the components contain chitin and chitosan, which have been recognized as potential sources materials for dietary supplement. This study aimed at evaluating the physical and chemical characteristics of chitosan nanoparticles extracted from sea cucumber *Holothuria scabra* when used as source materials for glucosamine. Chitin were extracted from dried samples, chitosan from chitin, while chitosan nanoparticles were obtained from chitosan with different concentrations (C_1 = 0.1%; C_2 = 0.2%; C_3 = 0.3%) of added sodium tripolyphosphate (NaTPP). Production process in this study resulted in 59.82% of chitosan extracted from chitin. Besides, the amount of chitosan nanoparticles obtained at 0.1%, 0.2%, and 0.3% additions of NaTPP were 90.6%, 92.8%, and 96.4%, respectively. These results were characterized in terms of whiteness degree (85.82%, 87.29%, 88.34%, respectively), deacetylation degree (90.6%, 95.8%, 96.2%), moisture (5.73%, 5.26%, 4.82%), and ash (1.29%, 1.07%, 0.98%). Looking at SEM and PSA tests, chitosan was morphologically found to be heterogeneously distributed with averaged 177-micron particle sizes. They also had larger particle chunks and solid as well as intact forms. Meanwhile, chitosan nanoparticles had smaller and smoother chunks, while they were produced in solid and intact forms. Besides, they were homogeneously distributed with sizes ranging between 134 – 206 nm (C_1), 114-128 nm (C_2), and 97-108 nm (C_3). Then, increments in NaTPP concentrations were discovered to contribute to the reduction of *H. scabra*-sourced chitosan nanoparticles size.

Keywords: Chitosan, *Holothuria scabra,* nanoparticles, sodium tripolyphosphate

ARTICLE INFO

Article history:
Received: 23 April 2019
Accepted: 19 July 2019
Published: 21 October 2019

E-mail addresses:
sumarto1976@yahoo.co.id (Sumarto)
bustarih@yahoo.com (Bustari Hasan)
kumans_69@yahoo.co.id (Rahman Karnila)
merysarmin@yahoo.com (Mery Sukmiwati)
* Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

INTRODUCTION

Sea cucumber is a marine animal originated from Indonesia. It has been recognized to offer various advantages in different aspects of human life, which may be either related or unrelated to dietary. In most situations, they are utilized as a functional food in health and biopharmaceutical fields, or as an ingredient in other chemical-related industries.

In particular, the commercial values of sea cucumber *H. scabra* J have been acknowledged due to its high nutrient contents, which in the form of flour offering a protein content of 60-70% with high essential amino acid components, containing complete fatty acids, carbohydrates, and minerals for human health. In the food industry, its uses begin with handling and weeding processes, in which parts of sea cucumber, particularly its layers that contain a lot of lime, are separated.

Technically, the anatomy of sea cucumber consists of 44.2% body mass (except viscera and gonads), 40% skins, 10.3% stomach contents (viscera and gonads), and 5.5% others including impurities. Despite having a similar proportion to general body parts, the skin has not been fully utilized yet, hence adding wasted parts of sea cucumber remains (Amri et al., 2018). Therefore, there is an opportunity to make it commercially available by increasing its economic values. Among others, the skin of sea cucumber may potentially be processed to produce chitin and chitosan, which is conducted by extracting the skin of sea cucumber to produce these two commercial products.

In general, the process of extracting chitosan was conducted by first soaking chitin shells/skins with NaOH 50% solution using a magnetic stirrer at 100 °C for 90, 120 and 150 minutes. The mix was then filtered and washed using distilled water to produce neutral pH. Then, chitosan being produced was dried in an oven at 45-47 °C for 6 hours. In the literature, extraction time that resulted in the most desired chitosan characteristics was 120 minutes, which produced chitosan in the form of a white fine powder with 37.59% yield, 6.48% water content, 8.19% ash content, and 69.29% deacetylation degree (Amri et al., 2018).

Practically, chitosan is applicable in various modern industries, including pharmacy, biochemistry, cosmetics, food and textile (Berger et al., 2004). The vast potentials of chitosan have been encouraging researches to expand the utilizations of chitosan products by modifying their chemical and physical properties. One of the physical modifications is applied by changing the size of its particles to smaller ones for wider utilizations, leading to the formation of nanoparticles. A smaller physical state is more advantageous compared to similar materials in a larger size due to a greater comparative value in terms of surface area and volume, making it more reactive. In fact, it fits with an established finding that refers to material reactivity as being determined by atoms on the surface, which come in a direct contact with other materials (Suwarda & Maarif, 2012).

In recent advances, chitosan nanoparticles have continued to be under investigation in terms of both determining their composition and finding an appropriate production method. In general, the production of high-quality chitosan nanoparticles requires a suitable, effective and simple method for obtaining a uniform size and desired stability. As a matter of facts, ionic gelation method has become a preferred method to obtain nanoparticles due to its simple process.

During an ionic gelation process, the formation of chitosan nanoparticles is conducted by reacting chitosan with sodium tripolyphosphate, which is known as a multivalent anion, to form cross-linked bonds with cationic chitosan. A method as such produces an interaction between the positive charge at chitosan's amino groups and the negative charge at tripolyphosphate, resulting in smaller particles (Lin et al., 2008). In fact, the production technology of chitosan nanoparticles offers critical advantage in producing good quality drugs by delivering desired characteristics and pellet size. Besides, chitosan is an important ingredient for producing glucosamine, therefore, producing the nanoparticles is expected to improve the quality of these products.

Looking at current literature, the production of chitosan nanoparticles by extracting chitosan shells from tiger shrimp through ionic gelation processes had been conducted by Nadia et al. (2014). In their study, the nanoparticle making process began with the mix of 0.1% Tripolyphosphate (TPP) solution with chitosan solution for then being homogenized by using a magnetic stirrer at 25°C temperature for 1 hour. The production of chitosan nanoparticles from tiger shrimp shells resulted in 80.67% yield rate, 98.65% deacetylation degree, 228.74 nm average particle size. In fact, the particle size is quite uniform, with a ball-like shape, and relatively stable.

In the literature, researches on chitosan nanoparticles from sea cucumber have not been conducted yet to utilize it as the primary ingredient for health-supplementing glucosamine production. Technically, glucosamine is a monomer of chitin and chitosan, which is often found in the shells/skins of various marine biota. Besides, glucosamine has been taken as food supplements to prevent and cure osteoarthritis. In human body, it is a precursor for the biosynthesis of glycosylate proteins and lipid to produce synovial fluid, which is used as a lubricant in the cartilage (Husskison, 2008).

Therefore, it is necessary to conduct a research on the production and quality characterization of chitosan nanoparticles by ionic gelation method using different concentrations of emulsifying solutions. This work hence aimed at determining the physical and chemical characteristics of chitosan nanoparticles extracted from the skins of *H. scabra* sea cucumber using different concentrations of sodium tripolyphosphate (NaTPP).

MATERIALS AND METHODS

Materials

As the focus of this research, several sea cucumbers (*H. scabra*) measuring 525±83.6 grams were obtained from the waters surrounding Terung Island, Batam, Indonesia. Besides, chemicals such as NaOH (Merck, German), acetic acid (Merck, German), distilled water (Bratachem, Indonesia), HCl (Merck, German), Sodium Tripolyphosphate (NaTPP; Merck, German), Tween 80 as emulsifier (Indonesia) and other ingredients were taken for proximate analyses. Equipment used for the preparation and handling of samples, the production of chitin and chitosan, and the manufacturing of chitosan nanoparticles included a magnetic stirrer (BIG LAB 79-1), oven, Particle Size Analyzer (PSA) (Beckman Coulter), Viscometer (Brookfield LV), Fourier Transform Infrared Spectrophotometer (FTIR) (MBQ00 Bruker Tensor Type), and Scanning Electron Microscopy (SEM) (JSM-35C).

Preparation and Flour-making Process

The dissection of sea cucumber being observed was done by using knife to split its abdomen, while fillets were used to separate gonad innards, other body mass and skin. The gonads, skin, other body mass and offal were dried separately in an oven at 40-45°C temperature for 48 hours. They were then grinded in a mixing machine (laboratory blender, model 32BL79, USA), which was set to produce 80 flour mesh size. Then, it was tapped repeatedly to ensure evenly grinded materials.

Chitin Extraction (No et al., 1989)

The sea cucumber flour was placed in a container. NaOH 3.4% was added to the container with 1:10 b/v ratio of ingredients and solutions. After that, the mixture was heated at 65°C for 2 hours while being continuously stirred. Next, the heated mixture was left to cool, then filtered and washed by using distilled water to reach a neutral pH. After the results were weighed, 1 N HCl was added into the tube with 1:10 b/v ratio of materials and solutions. The tube was re-heated at 65°C for 2 hours while being continuously stirred. Then, the precipitate was filtered and washed with distilled water to achieve a neutral pH, and dried at 60°C. The resulting product was chitin with determined chemical and physical characteristics.

Chitosan Extraction (Suptijah, 2004)

Furthermore, chitosan was obtained by making a highly concentrated solution of chitin deacetylation with 50% NaOH. In practice, chitin flour was weighed and 50% NaOH was added at 1:10 (b/v) ratio between chitin and solvent. The mixture was continuously stirred while being heated at 100°C for 120 minutes. Results of the deacetylation process were then deposited in a centrifuge for 15 minutes to separate solids from liquids. Obtained solids were repeatedly washed by using distilled water to achieve a neutral pH, and dried in an oven at 60°C for 6 hours.

Formation of Chitosan Nanoparticles (Iswandana et al., 2013)

Chitosan solution was produced by dissolving 200 mg of previously extracted chitosan in 100 ml of acetic acid 1% using a magnetic stirrer. The acetic acid 1% was obtained by mixing 10 ml of glacial acetate in 1000 ml of distilled. Next, a 0.1% concentration of sodium tripolyphosphate (NaTPP) solution was produced by dissolving 400 mg of NaTPP in 40ml of distilled and demineralized water by using a magnetic stirrer. Besides, a 0.2% concentration was made by dissolving 800 mg of NaTTP in 40 ml of distilled water, while a 0.3% concentration was produced by dissolving 1200 mg of NaTTP in 40 ml of distilled and demineralized water with a magnetic stirrer at 3000 rpm for 30 minutes to form a nanoparticle suspension. Next, the prepared chitosan solution was poured into a glass beaker while being stirred by using a magnetic stirrer. Then, separated NaTPP solutions (one for each concentration) was added slowly to the chitosan solution to form a nanoparticle suspension. The stirring was continued for 60 minutes to ensure a completed cross-linking process. In general, the whole process for making chitosan nanoparticles from sea cucumber took around 90 – 120 minutes.

Statistical Analysis

In this study, a statistical analysis on the results of experiments was conducted by applying a Completely Randomized Design (CRD). In practice, the data were analyzed by applying a one-way analysis of variance (ANOVA) in SPSS software version 22.0. Treatments being analyzed included NaTTP additions at different concentrations (C1 = 0.1 %; C2 = 0.2%; C3 = 0.3%) and a control chitosan (without any added NaTPP). Besides, test parameters covered moisture, ash, fat, protein, carbohydrate content (Association of Official Analytical Chemist – AOAC, 2005), yield, appearance, color, whiteness degree (Ernawati, 2012), the degree of acidity (Indonesian National Standards, 2004), the degree of deacetylation (Swann and Patwardhan, 2011), Scanning Electron Microscopy (SEM; Masooti et al., 2007) and Fourier Transform InfraRed (PSA; Yang et al., 2014).

Test Parameters

Moisture Content. This parameter was determined by applying the gravimetric method with an oven. The method involved weighing the moist sample after being dried in oven at 105 °C for 24 h. The water mass being produced was determined by comparing the weights of samples before and their constant weights after drying. Moisture content was calculated by using the following equation:

$$\% \text{ Moisture content} = \frac{\text{Wet weight (g)} - \text{Dry weight (g)}}{\text{Wet weight (g)}} \times 100$$

Ash Content. First, a tared crucible was dried and cooled. To discover the amount of ash in the prepared chitosan, 2 g of chitosan was placed into the tared crucible. Samples were heated in a muffle furnace at 600°C for 6 hours. The crucible was left in the furnace to naturally cool until reaching <200°C temperature, and then placed into a desicator for 30 minutes. Then, the mass of crucible and ash content was weighted.

Fat Content. Crude fat was determined by weighing 5 g of each sample to be wrapped in a filter paper by a Soxhlet apparatus using petroleum ether. It was conducted for 4 hours each. Next, extracted materials were left to evaporate all solvent content. After ensuring all solvent had evaporated, the extracted materials were weighed, and its fat content was calculated.

Crude Protein Content and Carbohydrate. Crude protein was analyzed by applying the Kjeldahl method (AOAC, 2005). Observed samples went through three essential steps, i.e. digestion, distillation, and titration, with 6.25 conversion factor to convert total nitrogen to crude protein. Thus, protein percentage in the samples could be calculated. Subtracting 100% by the sum of fat content, protein content, ash content, and moisture would then result in the total carbohydrate content (Onyeike et al., 2000).

Whiteness Degree (Ernawati, 2012). Determining the white degree of chitosan was conducted by using KETT Digital Whiteness Meter for Powder model C-100-3 (KETT Electric Laboratory, 1981). Samples were alternately put in measurement dishes until they were full and solid. Value indicated by the monitor referred to the white degree of observed sample (A), by which it was compared to a standardized whiteness value (110.8) according to the following equation:

$$\% \text{ WD} = \frac{A}{\text{Standard Value } BaSO_4 \text{ (110.8)}} \times 100$$

Degree of Deacetylation of Chitosan. The FTIR spectra of observed chitosan samples (in the forms of KBr disk and film) were obtained by using an I.R Instrument (MBQ00 Bruker Tensor Type) with *e* frequency range of 400-4,000 cm^{-1}. Deacetylation degree (DD) of the chitosan samples was calculated by following Khan et al. (2002) equation:

$$DD\% = 100 - \left[\left(\frac{A_{1655}}{A_{3450}}\right) x \left(\frac{100}{1,33}\right)\right]$$

Fourier Transform Infrared Spectroscopy (FTIR). Furthermore, the observed samples were characterized in an infrared spectroscopy by using KBr pellets with 400-4,000cm^{-1} scanning range (FTIR MBQ00 Bruker Tensor Type). KBr pellets were prepared (1 mg chitosan with 100 mg of KBr) and stabilized under a relative humidity before acquiring the spectrum (Brugnerotto et al., 2001). Next, transmittance or absorbance percentage was conducted by using an infrared spectrophotometry. Meanwhile, DD calculations of infrared spectrum in chitin and chitosan were conducted by comparing the absorbance of waves for NH-amide groups (1650-1500) cm-1 (A 1655) with that of primary amine group (3500-3200) cm-1 (A 3450) and the absorbance value of 1.33 for a perfect deacetylation process (Bastaman, 1989).

Scanning Electron Microscopy (SEM). Chitin skins, chitosan and its nanoparticles of *H. scabra* were examined by utilizing Scanning Electron Microscopy (SEM; Hitachi Flexsen 1000), which was equipped by EDS (Energy Dispersive X-Ray Spectroscopy) with two different magnification ranges (5000x for chitosan; 15000x for nanoparticles) and an accelerating voltage at 20kV (JSM-35C). It had a considerably large sample chamber and could accommodate samples as large as 300 mm in diameter and 110 mm high.

RESULTS AND DISCUSSION

Characteristics of Raw Materials

Characterization of raw sea cucumber was conducted to understand the proportion of each parts. Looking at the results of weeding on sea cucumber fillets, the discovered proportion included 40.10% skin, 43.53% other body mass, 10.97% stomach contents (viscera and gonads) and 5.40% impurities (Table 1, Figure 1). All proportions referred to their absolute comparison to total body weight.

Table 1
The average proportion of body parts for fresh sea cucumber (H. scabra J)

Parts of Raw Materials	Weight of Sea Cucumber* (g)	Fresh Proportion (%)	Weight of flour (g)	Yield of flour (%)
Skin	109.07	40.10	34.75	31.86
Other body mass	118.40	43.53	12.27	10.36
Stomach contents (Viscera and Gonad)	29.84	10.97	2.75	9.23
Impurities/leftovers	14.69	5.40	-	-
	272	100		

*frozen raw material

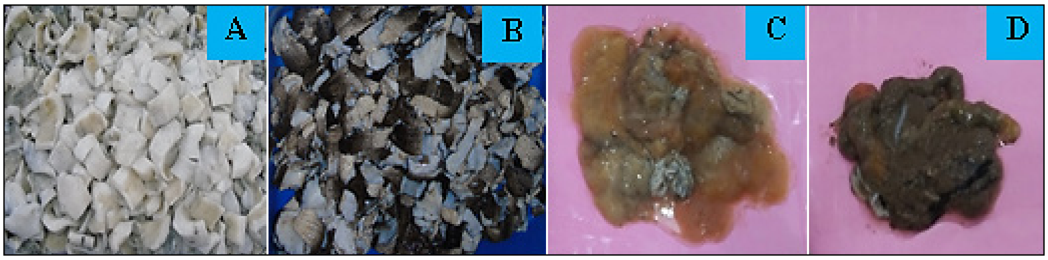


Figure 1. Body parts of sea cucumbers (A. body mass, B. skin, C. viscera and gonads, D. impurities/leftovers)

In fact, there were variations in the values of these characteristics before and after drying. For example, observations on the flour form discovered proportion of the skin at 31.86%, body mass at 10.36%, and stomach contents (viscera and gonads) at 9.23%. Significant differences were observed in the proportion of body mass and stomach contents (viscera and gonads) after drying and shaking processes due to their high level of moisture. Proximate analyses on the chemical composition of sea cucumber *H. scabra* revealed changes in composition between its fresh condition and after being transformed into flour form (Table 2).

Table 2
The chemical (proximate) content of the sea cucumber in fresh and flour raw materials

Chemical compositions	Composition of fresh sea cucumber and flour (%wb)					
	Skin		Other Body Mass		Viscera and Gonad	
	Fresh	Flour	Fresh	Flour	Fresh	Flour
Moisture	23.74 ± 0.52^a	8.25 ± 0.13^a	77.07 ± 0.51^b	9.12 ± 0.25^b	82.71 ± 0.22^c	16.12 ± 0.74^c
Protein	8.75 ± 0.34^a	12.12 ± 0.76^a	18.08 ± 0.48^c	72.25 ± 0.59^c	10.08 ± 0.19^b	43.47 ± 0.63^b
Fat	0.92 ± 0.05^a	0.64 ± 0.10^a	1.14 ± 0.11^b	1.95 ± 0.11^b	4.12 ± 0.26^c	16.85 ± 0.42^c
Ash	46.86 ± 0.27^c	55.03 ± 0.24^c	2.21 ± 0.14^a	5.76 ± 0.37^a	1.75 ± 0.19^b	11.98 ± 0.76^b
Carbohydrate	19.73 ± 0.51^c	23.95 ± 1.06^c	1.49 ± 0.34^b	10.62 ± 0.72^a	1.34 ± 0.40^a	11.58 ± 1.14^b

Numbers followed by same letters mean are not different really ($p<0.05$), data point are mean ± standard deviation (n=3)

In fresh condition, the moisture content was discovered to reach 23.74% (skin), 77.07% (body mass) and 82.71% (viscera and gonads). On the other hand, protein content was tested by following AOAC (2005) guidelines, resulting in 8.75% value (skin), 18.08% (other body mass), and 10.08% (viscera and gonads). Meanwhile, fat content was discovered according to the same guidelines, revealing 0.92% value (skin), 1.14% (other body mass), and 4.12% (viscera and gonads). Next, ash content was tested by applying the same guidelines, discovering 46.86% value (skin), 2.21% (other body mass), and 1.75% (viscera and gonads), while carbohydrate content (by-difference) amounted up to 19.73% (skin), 1.49% (other body mass), and 1.34% (viscera and gonads).

This study took the skin parts as raw materials for chitin and chitosan extraction processes. The results of proximate tests for the skin (Table 2), however, revealed variations in terms of its chemical composition between fresh condition and after being transformed into flour form. Technically, a change in form as such was aimed at facilitating the extraction process of chitin and chitosan. Changes occurring in the proportion of chemical composition due to the drying process (wet base) appeared to cause changes in the levels of water, protein, fat, ash and carbohydrates.

Characteristics of Chitin

Table 3 provides a comparison of the characteristics of chitin extracted from the flour form of sea cucumber to other sources. Looking at the results, a chitin yield of 39.08% was obtained from the skin of sea cucumber *H. scabra*. It was in fact relatively close to the result of prior work conducted by Amri et al. (2018) with 40.4% yield, while also higher than 27% yield obtained from crab shells (Nurjannah et al., 2016) and 33.24% obtained from snail shells (Dewi et al., 2016). However, yield discovered in this study was slightly lower to 45.08% yield from the extraction of *Portunus pelagicus* blue crab shell (Syukron et al., 2016). Furthermore, the characteristics of chitin produced by this study from the skin of sea cucumber *H. scabra* were found to mostly meet international standards (Table 3).

Table 3
Characteristics of chitin

Parameters	Chitin of sea cucumber	Chitin of shrimp[1]	Chitin of crab shells[2]	Chitin of blue crab shell[3]	Quality standard[4]
Yield (%)	39.08 ± 1.32	17.36	33.24	45.08	-
Moisture (%)	7.23 ± 0.71	8.50	-	5.72	≤ 10 %
Ash (%)	4.62 ± 0.33	4.25	-	4.84	≤ 2 %
Deacetylation degree (%) (FTIR)	38.84 ± 1.13	-	-	40.47	15-70 %

[1]Hossain and Iqbal (2014), [2]Dewi et al. (2016), [3]Syukron et al. (2016), [4]Bastaman (1989), Bastaman et al. (1990)

Furthermore, the purity of obtained chitin was observed by its low moisture content (7.23%) and DD (38.84%). In general, it was important to state DD as a parameter indicating the percentage of removable acetyl group from the deamination and deacetylation process. A high DD value indicated low acetyl group in the chitin. Technically, a reduction of the acetyl group would result in a stronger interaction between ions and hydrogen bonds (Winarti, 2008). On the other hand, ash content was found to be high (4.62%), exceeding the specified standard (≤ 2%). Then, the mineral content of sea cucumber skin was suspected to also be high, which was considerably not suitable for a demineralization process.

Characteristics of Chitosan and its Nanoparticles

Table 4 provides data on the characterization of chitosan and its nanoparticles. It showed the increasing concentrations of NaTPP from 0.1% to 0.3% to result in increased yield, color, whiteness degree, viscosity, and DD of chitosan nanoparticles. In fact, it was important to note the yield as being calculated from raw chitosan material, by which it was found to increase from 90.6% to 96.4% with respect to the increment of NaTTP concentrations. It appeared to be higher than 81.50% obtained from tiger shrimp shells (Nadia et al., 2014), 76.04% from green mussel shells (Suptijah et al., 2011), and 81.30% obtained from *Vannamei* shrimp shells (Arsyi et al., 2018). Furthermore, a longer stirring time was discovered to deliver a wider time frame for reducing particle sizes. In practices, the homogenization process between chitosan solution and ionic gelation material (NaTPP) could be controlled evenly at a high-speed set for a certain period of time, resulting in smaller chitosan particle sizes and relatively homogeneous particles.

Table 4
Characteristics of chitosan and its nanoparticles

Parameters	Chitosan (control)	C_1 (NaTPP 0.1%)	C_2 (NaTPP 0.2%)	C_3 (NaTPP 0.3%)	Quality standard
Yield (%)	56.84*	90.6**	95.8**	96.4**	-
Color	whiter	whiter	whiter	whiter	whiter
Whiteness degree (%)	80.27^a	85.82^b	87.29^b	88.34^b	-
Solution color (1.5%) (b/v)	clear	clear	clear	clear	clear
Moisture content (%db)	6.41^c	5.73^b	5.26^a	4.82^a	≤ 10 %
Ash content (%db)	1.41^b	1.29^b	1.07^a	0.98^a	≤ 2 %
Viscosity (cP) (1%)	426^c	198^b	176^a	162^a	Medium (200-799) Low (<200)
Deacetylation degree (%) (FTIR)	77.32^a	90.6^b	95.8^c	96.2^c	≤ 70 %

* yield of chitin material, ** yield of chitosan material

Furthermore, chitosan produced in this study was found to have characteristics that mostly met international standards (Table 4). Its purity could be observed from moisture and low ash contents despite a relatively near standard DD (≤ 70%). According to Suptijah (2006), a higher DD would result in more amine groups (NH_2) in chitosan molecule chains, making it increasingly reactive. Chitosan obtained in this work was in the form of granules, solid colloid, smooth, and whole round. According to Suptijah et al. (1992), on the other hand, their particle sizes were strongly influenced by raw materials being used. Chitosan derived in the current study from sea cucumber skin had a finer form, making it easy to get mashed up during chitosan production process (Table 5). Particle sizes as such also affected solubility, by which smaller particle sizes would make chitosan particles easier to dissolve in solvent.

Next, chitosan products in this study were found to be visually white in their powder form with slightly varied degrees of whiteness (80.27%-88.34%). The lowest whiteness degree was found for control chitosan (80.27%), while chitosan nanoparticles had white degrees up to 88.34% (Table 5). It hence emphasized the use of NaTPP concentrations to increase the whiteness degree of a nanoparticle product. Meanwhile, chitosan and chitosan nanoparticle products observed through color tests revealed a clear solution color in comparison to value standard. Prior work by Lisa et al. (2015) suggested color as one of critical parameters to determine the quality of flour products being produced. In general, consumers would prefer flour products with high white degrees. The whiteness degree for a flour product with a vastly brighter white color would reach as much as 100%.

After the solids of chitosan and chitosan nanoparticles were separately dried with 2-3mm thickness in an oven at 45 °C for 6 hours, the values of their water content were found to range from 4.82% to 6.41% (db). These values were in fact lower than the value standard for commercial chitosan products (≤10%). Technically, these values were affected by parameters applied in the drying process, including drying time, amount of dried chitosan, drying area, and drying techniques (Saleh et al., 1994).

The drying process reduced moisture in a material being dried through the evaporation of water during the heating process. Besides, changes occurred in terms of nutrient composition. For example, changes in the amount of ash could increase, while changes in product color might also occurr. In terms of product color, temperature being set should not be too high. The use of an excessively high heating temperature (> 60^{o}C) would damage the color of chitosan being produced, making it yellowish. Furthermore, ash content was a parameter used for determining minerals contained in chitosan, which might affect its solubility, viscosity, and characteristics of final product (No & Meyers, 1995). In this study, the levels of ash content obtained for chitosan and its nanoparticles being produced were in the range of 0.98%-1.41%, which appeared to fulfill specified quality requirements. In fact, it was important to note a low ash content as an indicator of low mineral content. Factors influencing the value might include the demineralization process and the washing technique, which used distilled water for pH neutralization (Angka and Suhartono, 2000). Theoretically, a good washing process would affect both ash and mineral levels released by a material being washed (Benjakula and Sophanodora, 1993).

Moreover, chitosan and its nanoparticles were found to have varying viscosities, *i.e.* 426 cP (control chitosan), 198 cP (0.1% NaTPP), 176 cP (0.2% NaTPP), and 162 cP (0.3% NaTPP). These values in fact met quality standards (200-700 cP in the medium category for raw/control chitosan; <200 cP in the low category for nanoparticles) suggested by Suptijah et al. (1992). The varied viscosities were considerably influenced by the deacetylation stage during production processes, in which the length of deacetylation process and high concentrations of NaOH would produce reduced molecular weight and viscosity. As the results, chitosan had a shorter chain compared to chitin, which was due to decreased molecular weight caused by the breakdown of polymeric bonds (depolymerization) of its molecular chains (Kolodziejska et al., 2000).

Then, DD value determined the amount of acetyl group lost during a deacetylation process. High DD values ranging from 77.32% to 96.2% indicated the purity of chitosan and nanoparticles being produced. Values as such, especially for chitosan nanoparticles, were found to meet quality standards (≥70%) suggested by Suptijah et al. (1992). These results slightly differed from 98.65% value obtained from tiger shrimp (Nadia et al., 2014). In their work, Muzarelli and Peter (1997) had stated a greater DD would result in a more active chitosan product, which was influenced by a large number of more reactive amine groups containing lone pairs of nitrogen atoms replacing acetyl groups in the structure.

Table 5
Morphological characteristics and the sizes of nanoparticles

Characteristics	Chitosan (raw, not nanoparticles)	C_1 (NaTPP 0.1%)	C_2 (NaTPP 0.2%)	C_3 (NaTPP 0.3%)	Quality standard
Particle shapes (SEM)	granules, solid colloid, smooth, and whole round	granules, solid colloid, smooth, and whole round	granules, solid colloid, smooth, and whole round	granules, solid colloid, smooth, and whole round	granules, solid colloid, and whole round
Particle sizes (PSA)	177^c micron	$134\text{-}206^b$ nm	$114\text{-}128^a$ nm	$97\text{-}108^a$ nm	10-1000 nm*
Yield (from raw chitosan,%)	-	90.6	92.8	96.4	≥80% **

*Mohanraj, (2006), **Suptijah et al. (2011)

In terms of morphology, Table 5 provides morphological characteristics of chitosan nanoparticles being produced. These nanoparticles could be visually distinguished by using SEM, which worked according to the nature of electron waves by applying diffractions at very small angles (Masooti et al., 2007).

Figure 2 exhibits the SEM testing of observed samples at different concentrations with magnifications up to 15,000x, while the testing on control chitosan applied up to 5,000x magnifications. Looking at the results, the shape of nanoparticles being observed was in the form of spheres resembling granules, solid colloid, smooth and whole round, and showed relatively homogeneous particle sizes.

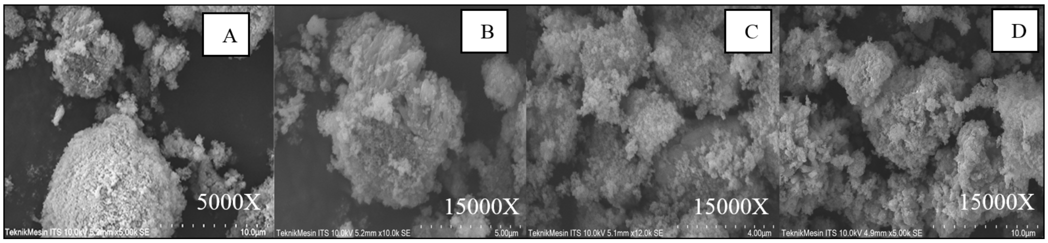


Figure 2. Morphology of chitosan (5,000x) and chitosan nanoparticles (15,000x) (A) Control of chitosan, (B) Nano chitosan 0.1% NaTPP, (C) Nano chitosan 0.2% NaTPP, (D) Nano chitosan 0.3% NaTPP.

In this study particle calculations were conducted by applying an image analysis. Chitosan nanoparticles being produced were tested by using a Particle Size Analyzer (PSA). The results showed average ranges of 134-206 nm (for 0.1% concentration), 114-128 nm (0.2%), and 97-108 nm (0.3%). In an agreement with these results, nanoparticles had been suggested to have solid-shaped particles with sizes ranging between 10-1,000 nm (Mohanraj, 2006). Technically, the method of preparing nanoparticles influenced their sizes. For example, the use of a magnetic stirrer would produce more stable particles with more even sizes under 1,000 nm (Al-Remawi, 2012). Besides, reducing particle sizes by utilizing a magnetic stirrer at a high-speed setting could spread energy received by all parts of a solution, making particle sizes increasingly homogeneous (Nesalin et al., 2009). Moreover, an appropriate NaTTP addition would produce reduced sizes of chitosan nanoparticles and an increased strength of chitosan matrix, making them stronger and harder to split (Du et al., 2009).

Then, obtained yields of chitosan nanoparticles were found to range between 90.6% to 96.4%, meeting the standard threshold (>80%). In fact, these values were greater than those of previous studies. According to Irianto and Muljanah (2011), magnetic stirrers offered advantages during the homogenization process between chitosan solution and NaTTP. Magnetic stirrers could be controlled evenly at high speeds to produce more homogeneous and stable particles with less to no agglomeration(s) for forming nanoparticles in the drying process. It was important to state the less-to-no agglomerated result as being applicable to stable particles only.

CONCLUSIONS

Chitosan produced from the skin of sea cucumber *Holothuria scabra* J offered relatively similar economic values compared to chitosan produced from other sources of raw materials (*e.g.*, shrimp shells and various crab types). This study revealed the use of different NaTTP concentrations to deliver various effects on the characteristics of chitosan nanoparticles obtained from the skin of *H. scabra* sea cucumber. It particularly affected the yield, whiteness degree, moisture content, ash content, viscosity, DD, and particle sizes of the nanoparticles. Product appearance and solution, on the other hand, were relatively similar, while their shapes were morphologically the same. Looking at the results of this study, the additions of NaTTP concentrations at 0.2% and 0.3% were discovered to deliver insignificant effects, resulting in relatively similar values of parameters being tested. Based on efficiency considerations, the 0.2% concentration of NaTPP was preferable. It was discovered to deliver optimal physical characteristics, resulting in 92.8% yield rate, a whiter color with 77.29% whiteness degree, clear solution color, desired particle shapes (granular, solid colloid, smooth and whole round), and smaller particle sizes ranging from 114-128 nm. The chemical characteristics included 5.26% moisture content (DW), 1.71% ash content (DW), and a low viscosity category (176 cP).

ACKNOWLEDGEMENT

This work was financially supported by the Society for Research and Service, the University of Riau, Pekanbaru, Indonesia, through a research funding granted during the 2018 fiscal year. All authors appreciate the helps, supports and efforts by everyone involved in the process of conducting this research.

REFERENCES

Al-Remawi, M. M. A. (2012). Properties of chitosan nanoparticlesformed using sulfate anions as crosslinking bridges. *American Journal of Applied Sciences, 9*(7), 1091-1100.

Amri, M. C., U., Ilza, M., & Sumarto. (2018). Characteristics of chitosan quality from skin sea cucumber *Holothuria scabra* with different heating time. *Journal of Priodic Fisheries of Terubuk, 46*(1), 1-10.

Angka, S. L., & Suhartono, M. T. (2000). *Utilization of marine waste: marine products biotechnology*. Bogor, Indonesia: Center for Coastal and Ocean Resources Studies, IPB University.

Arsyi, N. Z., Nurjannah, E., Ahlina, D. N., & Budiyati, E. (2018). Characterization of nano chitosan from green shells with ionic gelation method. *Natural Material Technology, 2*(2), 106-111.

Association of Official Analytical Chemist. (2005). *Official methods of analysis.* Washington, USA: Author.

Bastaman, S., Aprinita, N., & Hendarti. (1990). *Research on shrimp waste as an industrial material for chitin and chitosan*. Bogor, Indonesia: Center for Research and Development of Agricultural Products Industry.

Bastaman, S. (1989). *Studies on degradation and extraction of chitin and chitosan from prawn shells England*. Northern Ireland, UK: The Queen University of Belfast.

Benjakula, S., & Sophanodora, P. (1993). Chitosan production from carapace and shell of black tiger shrimp (*Penaeus monodon*). *Asean Food, 8*(4), 145-148.

Berger, J., Reist, M., Mayera, J. M., Feltb, O., Peppas, N. A., & Gurny, R. (2004). Structure and interaction in covalently and ionocally crosslinked chitosan hydragels for biomedical applications. *Europen of Pharm and Biopharm, 6*(57), 19-34.

Brugnerotto, J., Lizardi, J., Goycoolea, F. M., Argüelles-Monal, W., Desbrieres, J., & Rinaudo, M. (2001). An infrared investigation in relation with chitin and chitosan characterization. *Polymer, 42*(8), 3569-3580.

Dewi, N. L., Bahri, S., & Hardi, J. (2016). Utilizing Various of Pressures and Hidrolysis Times on Glucosamine Hydrochloride Production From Snail Shells (*Achatina fulica*) Chitosan. *Kovalen, 2*(1), 22-32.

Du , W. L., Niu, S. S., Xu, Y. L., Xu, Z. R., & Fan, C. L. (2009). Antibacterial activity of chitosan tripolyphosphate nanoparticles loaded with various metal ions. *Carbohydrate Polymers, 12*(75), 385-389.

Ernawati. (2012). *Making glucosamine hydrochloride (GlcNHCl) with the Autoclave Method*. Bogor, Indonesia: Department of Aquatic Product Technology, Faculty of Fisheries and Marine Science, IPB University.

Hossain, M. S., & Iqbal, A. (2014). Production and characterization of chitosan from shrimp waste. *Journal of the Bangladesh Agricultural University*, *12*(1), 153-160.

Huskisson, E. C. (2008). Glucosamine and chondroitin for osteoarthritis. *Journal of International Medical Research, 36*(6), 1161-1179.

Indonesian National Standards. (2004). *Water and Waste Water Part 11: How to Test the Acidity of pH using a pH Meter Tool*. Jakarta, Indonesia: National Standardization Agency.

Irianto, H. E, & Muljanah, I. (2011). The process and application of chitosan nanoparticles as drug carriers. *Squalen, 6*(1), 1-8.

Iswandana, R., Anwar, E., & Jufri, M. (2013). Formulation of Hydrochloride Verapamil Nanoparticles from Chitosan and Sodium Tripolyphosphate with Ionic Gelation Method. *Journal of Indonesian Pharmacy, 6*(4), 201-210.

KETT Electric Laboratory. (1981, Unpublished). Operating instruction kett digital whiteness-meter. Retrieved March 26, 2019, from https://www.kett.com/files/brc300.pdf.

Khan, T. A., Peh, K. K., & Ch'ng, H. S. (2002). Reporting degree of deacetylation values of chitosan: the influence of analytical methods. *Journal of Pharmacy and Pharmaceutical Science, 5*(3), 205-212.

Kolodziejska,I.,Wojtasz-Pajak,A.,Ogonowska, G., & Sikorski, Z. E. (2000) Deacetylationof chitin in a two-stage chemicaland enzymatic process. *Bulletin of the Sea Fisheries Institute, 2*(150), 15-24.

Lin, Y. H., Sonaje, K., Lin, K. M., Juang, J. H., Mi, F. L., Yang, H. W., & Sung, H. W. (2008). Multi-ion-crosslinked nanoparticles with pH-responsive characteristics for oral delivery of protein drugs. *Journal of Controlled Release, 132*(2), 141-149.

Lisa, M., Lutfi, M., & Susilo, B. (2015). Effect of temperature and duration of drying on quality flour of white oyster mushroom (*Plaerotus ostreatus*). *Journal of Tropical Agricultural Engineering and Biosystems, 3*(3), 270-279.

Masooti, A., Marino, F, Ortaggi, G., & Palocci, C. (2007). Fluorescence and scanning electron microscopy of chitosan/DNA nanoparticles for biological applications. *Journal Nanomedicine, 1*(4), 507-522.

Mohanraj, U. J. (2006). Nanoparticles-A review. *Tropical Journal of Pharmaceutical Research, 5*(1), 561-573.

Muzarelli, R. A. A., & Peter, M. G. (1997). *Chitosan Handbook*. New York, USA: European Chitin Society.

Nadia, L. M. H., Pipit, S., & Bustami, I. (2014). Production and characterization of nanocytosan from windu shrimp shell with ionic gelation methods. *Journal of Indonesian Fisheries Product Processing, 17*(2), 119-126.

Nesalin, A. J., Gowthamarajan, K., & Somashekhara, C. N. (2009). Formulation and evaluation of nanoparticles containing flutamide. *International Journal of ChemTech Research, 1*(4), 1331-1334.

No, H. K., & Meyers, S. P. (1995). Preparation and characterization of chitin and chitosan-A review. *Journal of Aquatic Food Product Technology, 4*(2), 27-52.

No, H. K., Meyer, S. P., & Lee, K. S. (1989). Isolation and characterization of chitin from crawfish shell waste. *Journal of Agricultural and Food Chemistry, 37*, 575-579.

Nurjannah, A., Darmanto, D., & Wijayanti, I. (2016). Optimization of the making of glucosamine hydrochloride (Glcn Hcl) from waste of crab shells through chemical hydrolysis. *Journal of Indonesian Fisheries Product Processing, 19*(1), 26-35.

Onyeike, E. N., Ayoologu, E. O., & Ibegbulam, C. O. (2000). Evaluation of the nutritional value of some crude oil in polluted freshwater fishes. *Global Journal of Pure and Applied Sciences, 6*(2), 227-233.

Saleh, M. R., Abdillah, S. E., Basmal, J., & Indriati, N. (1994). Effect of temperature, time and solvent concentration on chitosan extraction from frozen shrimp processing waste on several parameters of chitosan quality. *Journal of Post-Harvest Fisheries, 81*, 30-43.

Suptijah, P. (2004). Levels of chitosan quality modified results production process. *Agricultural Product Technology Bulletin, 7*(1), 56-67.

Suptijah, P. (2006). Characteristic description and application of chitin-chitosan. In *Proceedings of National Seminar about the Chitin Chitosan* (pp. 14-24). Bogor, Indonesia: Departement of Aquatic Product, Faculty of Fisheries and Marine Science, IPB University.

Suptijah, P., Jacoeb, A. M., & Rachmania, D. (2011). Characterization chitosan nano from shrimp shells (*Litopenaeus vannamei*) with ionic gelation methods. *Journal of Indonesian Fisheries Product Processing, 14*(2), 78-84.

Suptijah, P., Salamah, E., Sumaryanto, H., Purwaningsih, S., & Santoso, J. (1992). *Effect of Various Methods of Isolation of Chitin from Shrimp Skin to the Level and Quality* [Report]. Bogor, Indonesia: Faculty of Fisheries, IPB University.

Suwarda, R., & Maarif, M. S. (2012). Pengembangan Inovasi Teknologi Nanopartikel Berbasis PAT untuk Menciptakan Produk Yang Berdaya Saing [Innovative development of nanoparticle technology using PAT based to create a competitive product]. *Jurnal Teknik Industri, 3*(2), 104-122.

Swann, G. E. A., & Patwardhan, S. V. (2011). Application of fourier transform infrared spectroscopy (ftir) for assessing biogenic silica sample purity in geochemical analyses and palaeoenvironmental research. *Journal Climate of the Past, 10*(7), 65-74.

Syukron, F., Karnila, R., & Hasan, B. (2016). Characteristics of glucosamine hydrochloride (HCl GlcN) from chitin and chitosan swimming blue crab (*Portunus pelagicus*). *Journal of Priodic Fisheries of Terubuk*, 44(2), 22-35.

Winarti. (2008). Quality Characteristics and Solubility of Chitosan from Wind Shrimp Head Silage Dregs (*Penaeus monodon*). *Bulletin of Fishery Product Technology*. Retrieved March 13, 2019, from https://doi.org/10.17844/jphpi.v11i2.917.

Yang, M. H., Yuan, S. S., Huang, Y. F., Lin, P. C., Lu, C. Y., Chung, T. W., & Tyan, Y. C. (2014). A proteomic view to characterize the effect of chitosannanoparticle to hepatic cells. *Journal of BioMed Research International, 137*(6), 1-9.

Pertanika J. Sci. & Technol. 27 (4): 2427 - 2438 (2019)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Preparation and Characterization of Zeolite type 4A using Kaolin from Ajebo, Nigeria

Henry Ekene Mgbemere[1], Ikenna Christopher Ekpe[1]*, Ganiyu Lawal[1], Henry Ovri[2] and Anna-Lisa Chaudhary[3]

[1]Department of Metallurgical and Materials Engineering, University of Lagos, Akoka, Lagos State, Nigeria.
[2]Experimental Materials Mechanics, Institute of Materials Research, Helmholtz Zentrum Geesthacht Max Planck Strasse 1, 21502 Geesthacht Germany
[3]Department of Nanotechnology, Institute of Materials Research, Helmholtz Zentrum Geesthacht Max Planck Strasse 1, 21502 Geesthacht Germany

ABSTRACT

This work investigates the hydrothermal synthesis and characterization of zeolite-4A from kaolin found in Ajebo, Nigeria calcined at 700 and 900^{o}C respectively. The synthesized zeolite-4A was further characterised using X-ray Fluorescence (XRF), Fourier Transform Infrared spectrometer (FTIR), Scanning Electron Microscopy (SEM), X-Ray diffraction (XRD), Brunauer-Emmet-Teller (BET) surface area analysis as well as Differential Thermal Analysis/Thermo-gravimetric (TG). Water adsorption capacity tests were also carried out on the synthesized zeolite-4A. The results from the XRF measurements indicated that the amount of Al_2O_3 and SiO_2 in the studied kaolin was similar to the standard kaolin composition making it a perfect candidate for zeolite-4A synthesis. FTIR showed the characteristic zeolite peaks while XRD confirmed the crystalline nature of the synthesized zeolite-4A. TG studies showed that the zeolite-4A samples were stable up to temperatures of 700^{o}C. This stability as well as the surface area and pore size of 7 Å makes it potentially suitable for use in water treatment applications. The SEM showed cubic crystals which were typical of the morphology of zeolite-4A with water adsorption capacity of approximately 29%. These results indicate that zeolite-4A can be synthesized from kaolin found in Ajebo as an inexpensive alternative to traditionally sourced materials and also is suitable for use as adsorption agent.

Keywords: Adsorption agent, Ajebo kaolin, hydrothermal, metakaolin, zeolite-4A

ARTICLE INFO

Article history:
Received: 17 January 2019
Accepted: 16 April 2019
Published: 21 October 2019

E-mail addresses:
hmgbemere@unilag.edu.ng (Henry Ekene Mgbemere)
ekpeikenna@gmail.com (Ikenna Christopher Ekpe)
glawal@unilag.edu.ng (Ganiyu Lawal)
henry.ovri@hzg.de (Henry Ovri)
anna-lisa.Chaudhary@hzg.de (Anna-Lisa Chaudhary)
* Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

INTRODUCTION

Zeolites are commonly synthesized via hydrothermal methods with commercial chemicals as the primary sources of silica (SiO_2) and alumina (Al_2O_3) (Barrer, 1982). Chemicals such as sodium metasilicate, sodium aluminate, silica gel, tetraethylorthosilicate (TEOS) and aluminium hydroxide are used as precursor materials resulting in high purity, smaller pore sized zeolites. However, this pathway of production is complex and expensive thereby restricting the use of these materials in many industrial applications (Barrer, 1982; Breck, 1974; Chiang & Chao, 2001). As an example, a cost comparison of obtaining SiO_2 and Al_2O_3 sourced from a chemical feedstock of kaolin shows that by using kaolin as the precursor material, a cost advantage of about 15% over commercial chemical can be achieved (Vaculikova et al., 2011). The use of raw materials such as readily available and relatively cheap kaolin in zeolite synthesis is highly desirable as they contain SiO_2 and Al_2O_3 in the correct stoichiometry (Atta et al., 2007; Chandrasekhar, 1996; Chandrasekhar & Pramada, 1999; Yaping et al., 2008).

According to the report by Ugal et al. (2010) and the database of zeolite structures, zeolite type A or the Linde Type A (LTA), zeolites can be classified into three different grades, 3A, 4A and 5A, all of which possess the same general formula but with different structural cation types. When 75% of the Na^+ in type 4A is replaced by K^+, it is referred to as type 3A zeolite, alternatively replacing the Na^+ in type 4A by Ca^{2+} gives rise to zeolite type 5A.

Kaolin composition varies with location and the presence of impurities has an effect on the conversion to the final zeolite products (Bergaya et al., 2006). Therefore, it is important to source the material from a reliable region, hence, the significance of studying Ajebo kaolin as an inexpensive alternative with regards to the major zeolite properties such as adsorption.

The aim of this study was to synthesize and characterize zeolite-4A using kaolin from Ajebo as the precursor material and also to examine the effect of calcining at different temperatures.

MATERIALS AND METHODS

Sample Preparation

The base materials used in this research were kaolin from Ajebo (located in Ogun state, Nigeria with GPS coordinate 7.1229^o N, 3.6585^o E), NaOH and de-ionised water. The zeolite-4A was prepared according to the following procedure: The kaolin was initially crushed, ground and wet beneficiated to improve overall purity. After allowing the kaolin to soak for 24 hours, the top water and kaolin were separated by decantation. The slurry was sieved using a Tyler mesh sieve of mesh size 200 (75 μm aperture opening), allowed to settle and the supernatant water layer decanted in regular intervals until the slurry became thick. The thick slurry was then poured into a fabric sieve to drain off all the remaining liquid and pressure was applied until a cake like mass is formed. Finally, the solid cake was broken into smaller pieces and dried for 3 days after which it was milled in a ball milling machine made of steel cylinder, lined with ceramics to prevent contamination and loaded with ceramics ball with a ball to powder ratio (BPR) of 10:1 for 6 hours at a speed of 60 rpm. The calcination of the kaolin was done by controlled heating at a rate of 5^oC/min in a muffle furnace at 700 and 900^oC respectively and held at these temperatures for 2 hours followed by cooling in air. This process helps in the conversion of the unreactive and highly stable kaolin to a more reactive and less stable metakaolin which easily accepts and exchanges ions in its lattice.

The metakaolin was mixed with NaOH and deionised water in the ratio 0.6:0.8:12. The mixture was aged for 24 hours with intermittent mechanical stirring. Crystallization was achieved by heating the mixture in an oven operating at a temperature of 100^oC for 6 hours. After crystallization the synthesized zeolite-4A was washed three times with distilled water to remove any excess unreacted NaOH. The synthesized zeolite-4A was placed into a centrifuge rotating at a speed of 20 rpm for 20 minutes to separate the zeolite from the water. After the water has been decanted, the residual zeolite-4A was dried in an oven at 110^oC for 6 hours.

Sample Characterization

The composition of the kaolin and the as-synthesized zeolite-4A were analysed using an XRF (AXIOS, PANalytical, Netherlands). The samples were pressed into 6 mm thickness pellets using wax as a binder with a pellet to wax ratio of 11:3. The thermal analyses (Differential Thermal Analysis/Thermogravimetric (DTA/TG, Netzsch STA 409 C, Germany) were carried out in 50 mL/min argon flow from 20 to 800^oC at a heating and cooling rate of 5^oC/min respectively.

Synchrotron X-Ray diffraction measurements were done at the Deutsches Elektronen Synchrontron (DESY) radiation facility in Hamburg at the beamline P02.1 of PETRA III at a photon energy of 60 KeV (λ= 0.20716Å). The data were collected with a Perkins Elmer 2D detector and radially integrated in 5^o steps (Dippel et al., 2015; Herklotz et al., 2013). The Fourier Transform Infrared spectroscopy was carried out with a Schimadzu IR Affinity-1S (Schimadzu Corporation, Japan). The powder samples were pressed into pellets by mixing with dried KBr and placing them in a pre-heated die under a heat lamp. The pellets were degassed at 120^oC for 1 hour prior to obtaining the spectra. The spectra typically had an average of 128 scans with 1 cm^{-1} resolution. The infrared spectra were acquired in transmission mode while the recorded peaks were based on percentage transmittance to the given wavelengths.

The morphology of the samples was acquired using SEM (ASPEX 3020, Aspex Corporation, USA). The samples were coated with 10 nm of gold to inhibit charging and improve the secondary electron signals. The Brunauer-Emmett-Teller (BET) nitrogen adsorption and surface area analysis were carried out using a Micromeritics ASAP 2020 porosity analyzer (USA). The samples were initially degassed at 80°C under vacuum for 40 hours while the cold and warm free space values were determined using Helium. The samples were subsequently degassed again at 350°C for 8 hours under vacuum and N_2 sorption determined using the previously measured free space values. The water adsorption capacity of the zeolite-4A was measured using a custom-built apparatus where high humidity air flow was introduced to a sample of the zeolite-4A. The increase in weight of the zeolite-4A with adsorbed water was then measured after 24 hours.

RESULTS AND DISCUSSION

Composition Analysis

The chemical composition of the kaolin and the synthesized zeolite-4A is shown in Table 1. The major constituents in the kaolin are SiO_2 and Al_2O_3 representing 51.38 wt.% and 42.05 wt.% respectively with a SiO_2/Al_2O_3 ratio of approximately 1. The amount of Fe_2O_3 in the kaolin is less than 1 wt.% making it suitable for use in zeolite synthesis. Since the XRF of the zeolite-4A synthesized from kaolin calcined at 700 and 900°C is similar, the XRF of the zeolite-4A synthesized from kaolin calcined at 900°C was presented in Table 1. The amounts of Na_2O in the zeolite increased from 0.75 to 3.42 wt.%. This increase in the amount of Na_2O is due to the introduction of NaOH during the zeolite-4A synthesis. While all components present in the kaolin can been seen in the synthesized zeolite-4A indicating that the kaolin is stable at temperatures below 900°C.

Table 1
The chemical composition of Ajebo kaolin and Zeolite-4A

	Compound Amounts (weight %)									
	SiO_2	Al_2O_3	Fe_2O_3	CaO	MgO	TiO_2	Na_2O	K_2O	MnO	L.O.I
Kaolin	46.48	39.81	0.79	0.16	0.034	1.99	0.75	0.3	0.001	9.68
Zeolite-4A	58.9	22.76	0.15	0.05	0.05	0.019	3.42	0.89	0.005	13.756

Fourier Transform Infrared Spectroscopy (FTIR)

The FTIR spectra of the zeolite-4A synthesized from kaolin calcined at 700 and 900°C are shown in Figure 1. The FTIR spectra of the kaolin and metakaolin have been discussed elsewhere (Mgbemere et al., 2018). Both spectra show similar wavenumbers with a well-defined peak at 1000 cm^{-1} and 1655 cm^{-1}. There are differences in wavenumbers between the zeolite-4A calcined at 700°C and 900°C. This includes the absence of peaks at wavenumbers, 2200 cm^{-1} and 1500 cm^{-1} for the 900^{o}C sample and differences in the peaks below wavenumbers 800 cm^{-1}. It has been enlarged to the right to highlight the differences which are associated with the degree of conversion from the metakaolin to zeolite-4A. The samples calcined at 900^{o}C converted better than those at 700^{o}C. The most significant vibration at 1000 cm^{-1} region is assigned to a T-O (where T can be an Al or a Si ion) stretching which involve motions that are primarily either associated with oxygen atoms or described as an asymmetric stretching mode $O - T \leftrightarrow O$ (Mozgawa et al., 2005).

The weak peaks below 1000 cm^{-1} wavenumber are possibly due to the vibration involving the ≡Al-OH nests created by the cation vacancies. The peaks at 700-750 cm^{-1} corresponds to the symmetric stretching vibration of SiO_4 groups while the bands around 450-650 cm^{-1} are related to either bending vibration of SiO_4 groups or in the vibration modes of the 4-membered rings of the silicate chains. The splitting of the peak at this region which is possibly due to higher O-Si-O angles in the structure (Kovo, 2011). The weak broad band at 3700-4000 cm^{-1} is attributed to the stretching vibration of both hydrogen group of H_2O molecules and amines present in the pores of the zeolites. After calcination, H_2O is reintroduced during the process of zeolite formation. The peaks between 1500 and 1700 cm^{-1} are from the bending vibration of H_2O while those at 3300 cm^{-1} and 3500 cm^{-1} are as a result of the OH stretching modes (Mozgawa et al., 2005).

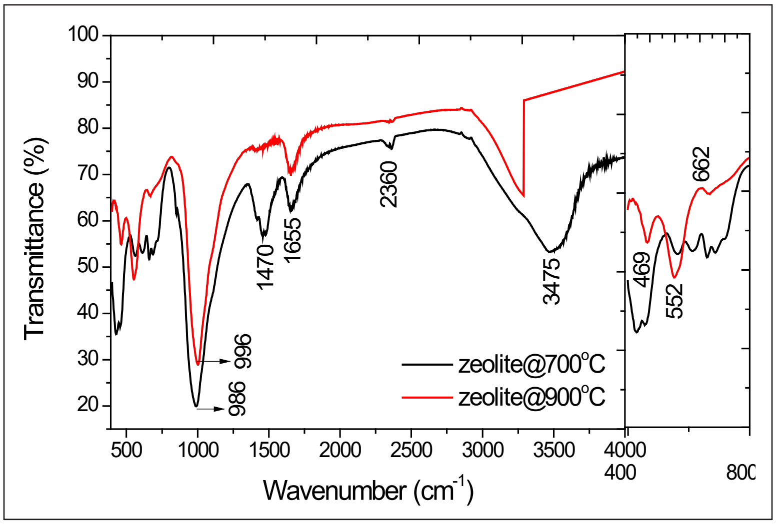


Figure 1. A graph of Transmission as a function of wavenumber for zeolites-4A samples from kaolin calcined at 700^{o}C and 900^{o}C. The inset on the right is the enlargement of the wavenumbers from 400 to 800 cm^{-1}

X-ray Diffraction

The XRD patterns for the synthesized zeolite-4A from kaolin calcined at 700 and 900^{o}C measured under normal conditions are shown in Figure 2. Both heat treated samples show Bragg angles correlated to the formation of zeolite-4A as seen in 74-1183-*ICDD, 2001*

(Baerlocher et al., 2007; Ugal et al., 2010) and confirmed by the Structural Commission of the International Zeolite Association (Treacy & Higgins, 2001). The difference between the zeolites produced from calcining at both temperatures is in the intensity of their peaks as the zeolite calcined at 900°C has slightly more intense peaks indicating that the zeolite-4A synthesized from kaolin calcined at 900°C is more crystalline than that synthesized from kaolin calcined at 700°C. Zeolite-4A is the major constituent phase while the impurity phases are from quartz, hydroxysodalite and possibly unconverted metakaolin. The extra peaks observed in the patterns are due to the presence of other oxides associated with the precusor kaolin material.

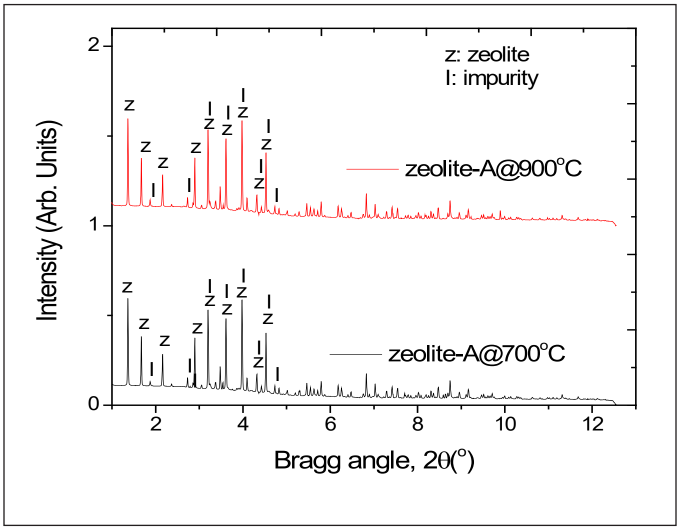


Figure 2. High resolution x-ray diffraction (λ= 0.20716Å) patterns for zeolite-4A samples obtained from kaolin calcined at 700°C and 900°C respectively

Thermal Analysis of Zeolite-4A

The TG/DTA graphs of the zeolite-4A synthesized from kaolin calcined at 700 and 900°C are presented in Figure 3. Both zeolite samples exhibit very comparable trends in the TG/DTA measurements. The mass loss during the TG measurements is approximately 13.67% for both zeolite-4A samples. This value is similar to that of the loss on ignition (LOI) from the chemical analysis in Table 1. The DTA graph is used to show the reactions (exothermic and endothermic) that took place as samples were heated from room temperature. Strong exothermic reactions are observed at 243.09°C and 247.56°C for the zeolite synthesized from kaolin calcined at 700°C and 900°C respectively. The absence of a loss at temperatures below 100°C may be due to the absence of physically adsorbed moisture as the zeolites were effectively dried after synthesis. The DTA curves for the zeolite-4A indicate that the zeolites are stable at temperatures up to 700°C.

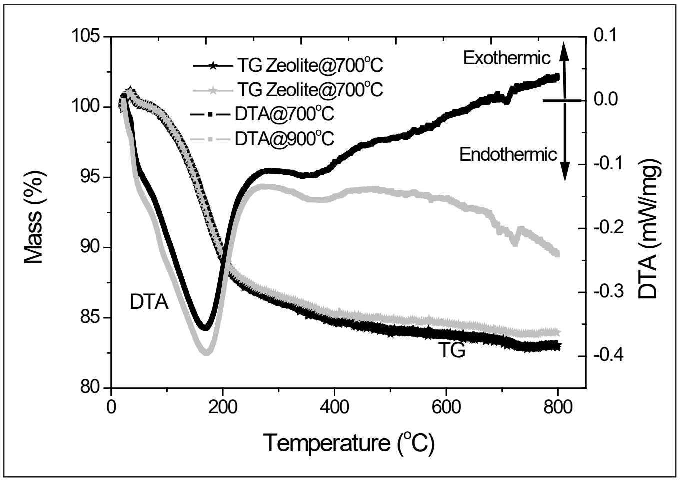
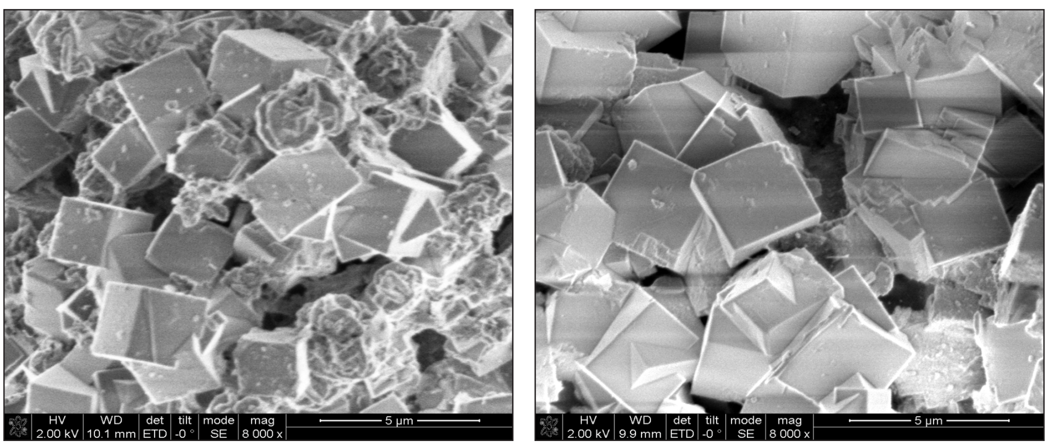


Figure 3. TG/DTA curves for Ajebo zeolite-4A calcined at 700°C and 900°C respectively. The data was acquired on heating the sample from room temperature

Scanning Electron Microscopy

The SEM image of the zeolite-4A synthesized from kaolin calcined at 700 and 900°C are presented in Figure 4. The morphology of the zeolite-4A samples shows cubic crystalline structures with well-defined edges with average sizes of 5 μm as well as aggregates especially for zeolite-4A calcined at 700°C. The EDS of the aggregates show that the major elements Al, Si, O and C are present. The starting materials for the synthesis of zeolites have an effect on the size of the zeolite-A crystals produced. Naturally occurring raw materials such as kaolin and fly ash tend to form larger zeolite crystals due to the presence of other oxides such as F_2O_3, CaO, and MgO in the kaolin (Petkowicz et al., 2008). These



(a) (b)

Figure 4. Scanning electron microscope images for Zeolite-4A synthesized from kaolin calcined at (a) 700°C and (b) 900°C

oxides serve as heterogeneous nucleation sites for the crystals and also help to quicken the crystallization kinetics. The nature of the metakaolin used to synthesize the zeolite can also determine the shape of the crystals. For instance for kaolin that has not completely undergone the meta-kaolinization process during the synthesis of zeolite-4A, Sodalite is produced instead of Zeolite-A (Reyes et al., 2010; Zhao et al., 2010). The Si/Al ratio also appears to have an effect on the shape of the crystals. For example, when the Si/Al ratio is 1.5, beveled edges as opposed to sharp well defined edges are obtained (Ismail et al., 2010).

Pore Size Analysis

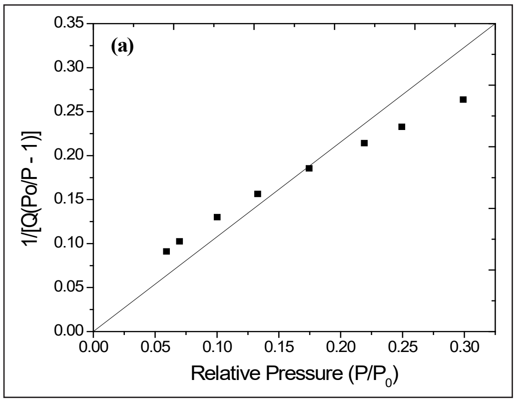
The BET surface area, Langmuir surface area, pore volume and pore size of both samples have been summarized in Table 2. Both zeolites have lower surface areas due to the larger crystal size described in the previous section. Moreover, the surface area of each sample decreased with an increase in calcination temperature used during the process of zeolite synthesis. Higher calcination temperature results in a decrease in surface area and this can be attributed to the further crumbling and shrinkage of the metakaolin structure at high temperatures. The pore volume and pore size increases with an increase in calcination temperature hence, the sample calcined at a higher temperature (900^{o}C) appears to have a looser amorphous structure than that calcined at lower temperature (700^{o}C). The pore size of the zeolite-4A synthesized from Ajebo kaolin is higher than that reported for type A zeolite which is approximately 4Å (Mgbemere et al, 2017; Ugal et al., 2010) when pure Al_2O_3 and SiO_2 are used. The difference in pore size between the zeolite-4A synthesized from pure chemicals and that synthesized from kaolin may be due to the presence of oxides other than SiO_2 and Al_2O_3 in the kaolin. As stated earlier, these impurities serve as nucleation sites for the formation of zeolite crystals hence increasing the pore size of the zeolite (Farag & Zhang, 2012).

Table 2
Brunauer-Emmet-Teller (BET) analysis of zeolites calcined at 700^{o}C and 900^{o}C respectively

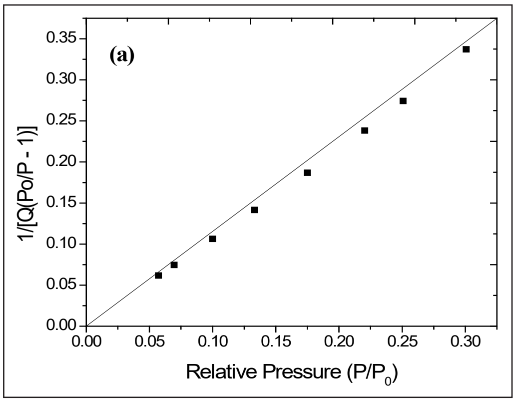
Compound Amounts (weight %)		
	Ajebo 700^{o}C	Ajebo 900^{o}C
BET Surface area (m^2/g)	5.6629	5.5472
Langmuir Surface Area (m^2/g)	5.4165	5.0353
Pore volume (cm^3/g)	0.009897	0.002000
Pore size (nm)	6.999103	7.58527

Figure 5 shows the BET surface area plot for both zeolite-4A synthesized from kaolin. The changes in surface area and crystallinity have been measured using the BET N_2 adsorption. According to Yates (Yates, 1967), for fresh zeolites, all points falls on a 45^{o} line with no loss in surface area or crystallinity but when majority of the points falls above the line, it indicates less loss in crystallinity than in surface area while when majority of the points falls below the line it indicate more loss in surface area than in crystallinity.

From figure 5a below, there are 4 points below the 45^o line and 4 above it which is an indication that there is an equal amount of losses in the surface area as well as in crystallinity of both zeolites. While from figure 5b all point falls below the 45^o line an indication that there are greater lose in surface area than in crystallinity.



(a)



(b)

Figure 5. BET surface area plot of zeolite-4a samples from kaolin calcined at (a) 700^{o}C and (b) 900^{o}C

Figure 6 shows the nitrogen adsorption/desorption isotherm for the synthetic zeolite-4A. The location of the hysteresis loop in the N_2 isotherm is a type IV hysteresis loop which is an indication of the presence of mesoporosity. This is used to determine the type of pores present in the zeolite for example, a regular framework pore or inter-particle voids such as

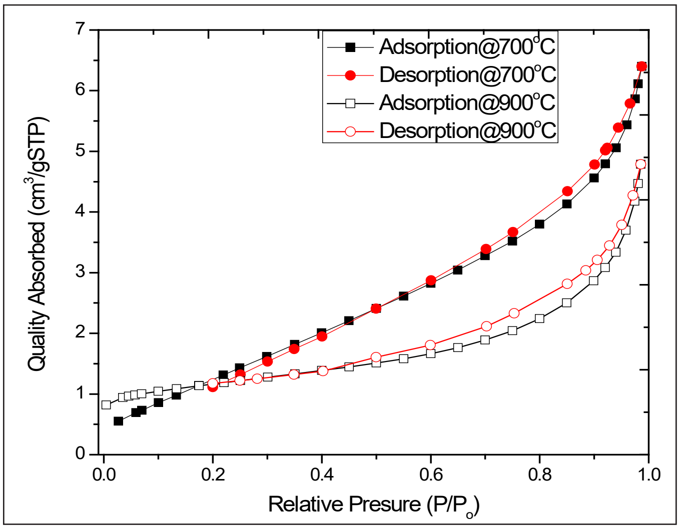


Figure 6. N_2 adsorption/desorption isotherms for the synthesized zeolite-4A samples

a textured pore. The porosity between 0.4 and 0.75 P/P_o in the N_2 isotherm of the zeolite-4A gives an indication that the porosity is related to a framework porosity and a value between 0.8 and 1 P/P_o is a textural porosity arising from aggregate voids and spaces formed by intra-particle contacts. The zeolite-4A synthesized from kaolin calcined at 700°C shows a higher rate of adsorption and desorption compared to the zeolite-4A synthesized from kaolin calcined at 900°C, this is as a result of the presence of more amorphous gel in the zeolite-4A synthesized from kaolin calcined at 700°C.

Water Adsorption Capacity Test

Table 3 shows the water adsorption capacity data for zeolite-4A synthesized from kaolin calcined at 700 and 900°C. The synthesized zeolite-4A shows highly promising prospects for water adsorption as the water adsorption capacity was found to be 28.89 and 28.96 % for the zeolite-4A synthesized from kaolin calcined at 700 and 900°C respectively. These are higher than the 25.9 % for commercial zeolite-A reported in the literature (Miao et al., 2009).

Table 3
Water adsorption capacity results for Zeolite-4A synthesized from Ajebo kaolin

Water Adsorption Capacity Test		
	zeolite@700°C	zeolite@900°C
Wt. of Beakers (g)	104.48	48.64
Wt. of Beaker + Wt. of Dry Zeolite-4A (g)	108.12	52.21
Wt. of Beaker + Wt. of wet Zeolite-4A (g)	109.17	53.24
% Water Adsorption	28.89	28.96

CONCLUSION

Crystalline, microporous zeolite-4A was successfully synthesized using Ajebo kaolin as the precursor material. Temperatures of 700 and 900°C were used successfully to convert kaolin to metakaolin for the purpose of synthesizing the zeolite-4A. The DTA/TGA plots showed that both zeolites were stable at temperatures below 700°C with a weight loss of about 13.67 wt.% due to loss of water of crystallization. The FTIR spectra and XRD patterns confirmed the presence of characteristic zeolite peaks with the zeolite-4A synthesized from kaolin calcined at 900°C having higher peaks and intensity indicating better crystallinity than that synthesized from kaolin calcined at 700°C. While the SEM clearly showed the formation of cubic crystal structured zeolite-4A at both calcination temperatures with that synthesized from kaolin calcined at 900°C resulting in more crystals and less amorphous gel. The synthesized zeolite-4A showed promising signs for use in water adsorption, as both had water adsorption capacity of 28.89 and 28.96 % for the zeolite-4A synthesized from kaolin calcined at 700 and 900°C with that from 900°C being slightly higher. The zeolite-4A synthesized from Kaolin calcined at 900^{o}C showed better properties than that synthesized from kaolin calcined at 700^{o}C.

ACKNOWLEDGEMENT

The authors acknowledge the financial support from the Central Research Committee of the University of Lagos under the grant: CRC 2016:14. The authors are grateful to Dr. Mauricio Schieda of the Department of Sustainable Energy Technology Institute of Materials Technology at the Helmholtz-Zentrum Geesthacht for access to and support for the BET measurements and Dr. Manuel Hinterstein for the synchrontron diffraction measurements.

REFERENCES

Atta, A. Y., Ajayi, O. A., & Adefila, S. S. (2007). Synthesis of faujusite zeolites from kankara kaolin clay. *Journal of Applied Science Research, 3*(10), 1017-1021.

Baerlocher, C., McCusker, L. B., & Olson, D. H. (2007). *Atlas of zeolite framework types*. Amsterdam: Elsevier.

Barrer, R. M. (1982). *Hydrothermal chemistry of zeolite*. New York: Academic Press.

Bergaya, F., Theng, B. K. G., & Lagaly, G. (2006). *Handbook of clay science*. Amsterdam: Elsevier.

Breck, D. W. (1974). *Zeolite Molecular Sieve: Structure, Chemistry and Use*. New York: John Wiley & Sons Inc.

Chandrasekhar, S. (1996). Influence of metakaolinization temperature on the formation of zeolite 4A from kaolin. *Clay Mineral, 31*(2), 253-261.

Chandrasekhar, S., & Pramada, P. N. (1999). Investigation on the synthesis of zeolite NaX from Kerala kaolin. *Journal of Porous Materials, 6*(4), 283-297.

Chiang, A. S. T., & Chao, K. J. (2001). Membranes and films of zeolite and zeolite like materials. *Journal of Physics and Chemistry of Solids, 62*(9-10), 1899-1910.

Dippel, A. C., Liermann, H. P., Delitz, J. T., Walter, P., Schulte-Schrepping, H., Seeck, O. H., & Franz, H. (2015). Beamline P02.1 at PETRA III for high resolution and high energy powder diffraction. *Journal of Synchrotron Radiation, 22*(3), 675-687.

Farag, I. H., & Zhang, J. (2012). Simulation of synthetic zeolites-4A and 5A, manufacturing for green processing, IRACST. *Engineering Science and Technology: An International Journal, 2*(2), 188-195.

Herklotz, M., Scheiba, F., Hinterstein, M., Nikolowski, K., Knapp, M., Dippel, A. C., … & Ehrenberg, H. (2013). Advances in in situ powder diffraction of battery materials: a case study of the new beamline PO2.1 at DESY, Hamburg. *Journal of Applied Crystals, 46*(1), 1117-1127.

Ismail, A. A., Mohamed, R. M., Ibrahim, I. A., Kini, G., & Koopman, B. (2010). Synthesis, optimization and characterization of zeolite and its ion exchange properties. *Colloids and Surfaces: Physicochemical Engineering Aspects, 36*(6), 80-87.

Kovo, A. S. (2011). *Development of Zeolites and zeolite membranes from Ahoko Nigerian Kaolin*. (Doctoral dissertation). The University of Manchester, United Kingdom.

Mgbemere, E. H., Ekpe, C. I., & Lawal, I. G. (2017). Zeolite synthesis, characterization and application areas: A review. *International Research Journal of Environmental Sciences, 6*(10), 45-59.

Mgbemere, H. E., Lawal, G. I., Ekpe, I. C., & Chaudhary, A. L. (2018). Synthesis of zeolite-a using kaolin samples from Darazo, Bauchi State and Ajebo, Ogun State in Nigeria. *Nigerian Journal of Technology, 37*(1), 87-95.

Miao, Q. Z., Zhou, Z., Yang, J., Lu, S., & Wang, J. (2009). Synthesis of NaA zeolite from kaolin source. *Frontiers of Chemical Engineering in China, 3*(1), 8-11.

Mozgawa, W., Jastrzebski, W., & Handke, M. (2005). Vibrational spectra of D4R and D6R structural units. *Journal of Molecular Structure, 5*, 663-670.

Petkowicz, D. I., Rigo, R. T., Radtke, C., Pergher, S. B., & Santos, J. H. (2008). Synthesis of zeolite NaA from Brazillian chrysotile and rice-husk. *Microporous and Mesoporous Materials, 1*(16), 548-554.

Reyes, C. D., Williams, O., Mauricio, R., & Alarcon, C. (2010). Synthesis of zeolite LTA from thermally treated kaolinite. *Revista Facultad De Ingenieria, University of Antioquia, 53*, 30-41.

Treacy, M. M., & Higgins, J. B. (Eds.). (2001). *Collection of simulated XRD powder patterns for zeolites* (4th Ed.). Amsterdam: Elsevier.

Ugal, J. R., Hassan, K. H., & Ali, I. H. (2010). Preparation of type 4A zeolite from Iraqi kaolin: characterization and properties measurement. *Journal of the Association of Arab University for Basic and Applied Science, 9*(1), 2-5.

Vaculikova, L., Plevova, E., Vallova, S., & Koutnik, I. (2011). Characterization and differentiayion of kaolinites from selected Czech deposites using infrared spectroscopy and differential thermal analysis. *Acta Geodyn, Geomater, 8*(1), 59-67.

Yaping, Z., Xiaoqiang, Q., Weilan, Y., & Mingwen, W. (2008). Synthesis of Pure zeolite from supersaturated silicon and aluminium alkali extract from coal ash. *Fuel Processing Technology, 8*(7), 1880-1886.

Yates, D. J. C. (1967). Studies on the surface area of zeolite, as determined by physical adsorption and X-ray crystallography. *Cadadian Journal of Chemistry, 46*(10), 1695-1701.

Zhao, Y., Zhang, B., Zhang, X., Wang, J., Liu, J., & Chen, R. (2010). Preparation of highly ordered cubic NaA zeolite from halloysite mineral for adsorption of ammonium ions. *Journal of Hazardous Materials, 17*(8), 658-664.

Pertanika J. Sci. & Technol. 27 (4): 2439 - 2452 (2019)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Scattered Radiation Dose Calculation in the Multi-Bay Resuscitation Room using MCNP5 Code

Marianie Musarudin[1]*, Ruzainah Mohd Nordin[1] and Nik Kamarullah Ya Ali[2]

[1]School of Health Sciences, Health Campus, Universiti Sains Malaysia, 16150 USM, Kubang Kerian, Kelantan, Malaysia

[2]Department of Radiology, Hospital Universiti Sains Malaysia, 16150 USM, Kubang Kerian, Kelantan, Malaysia

ABSTRACT

This study aims to characterise the scattered dose distribution from a ceiling-mounted X-ray unit in a multi-bay resuscitation room. The finding of this study is essential for optimisation and safety of staff and patients. Simulation of phantom imaging was carried out using MCNP5 code. The calculated data were initially compared against the measurements carried out using a survey meter. Three measurement positions, denoted by T2, T3, and T4 were considered for the dose calculation. The data suggested that T2 received the highest scattered dose. This value (maximum value of less than 6 μGy) is lower than the annual dose limit for the public and radiation workers as well as natural background radiation dose. Meanwhile, T3 consistently received a higher scattered dose (maximum difference of 25.62%) than T4. The angles of the X-ray tube resulted in scattered doses less than 6 μGy for both 90^{o} and 100^{o} scattering angles. In conclusion, the scattered dose for a single exposure imaging inside the room is safe. Yet, consideration of the placement of a portable lead shielding between X-ray tube and treatment couch is strongly recommended. This is due to a high number of imaging procedures commonly performed daily in a busy hospital. Hence, the cumulative dose to the paramedic staff and patients may exceed the safe level.

Keywords: MCNP5, scattered dose, x-ray room

ARTICLE INFO

Article history:
Received: 13 May 2019
Accepted: 03 July 2019
Published: 21 October 2019

E-mail addresses:
marianie@usm.my (Marianie Musarudin)
eina8_nordin@yahoo.com.my (Ruzainah Mohd Nordin)
kamarullah@usm.my (Nik Kamarullah Ya Ali)
* Corresponding author

INTRODUCTION

The guidelines for a safe level of exposure to radiation for medical purposes for staff and patient have been tightened by the International Commission on Radiological Protection (ICRP). This enforcement is due to the increase in the use of radiation for

ISSN: 0128-7680
e-ISSN: 2231-8526

medical purposes. In comparison to the early 1980s, the exposure to ionizing radiation from medical diagnostics has increased up to seven times. According to Report No. 160 released by the National Committee on Radiological Protection (NCRP) in 2009, the collective effective dose for X-ray examinations represented 5% of all exposure categories (David & Otha, 2009).

Measurement of the doses for an X-ray room is commonly performed at several points outside the planned location for a shielding barrier. For instance, doses at the walls and doors are commonly assessed as a part of radiation protection monitoring. This is an important practice to determine the amount of scattering inside and outside the radiation room. The primary, scatter, and leakage radiations are necessary factors to be considered during the design phase of the X-ray room. Therefore, previous studies had commonly presented the transmission data for various materials of the shield used in X-ray room such as lead, concrete, gypsum wallboard, steel, and plate glass (Archer et al., 1994; Simpkin, 1989; Simpkin, 1995). The difference between the penetrating powers of the X-ray produced by single and three-phase generators has also been highlighted in these studies. In other studies, the main focus was on the amount of dose absorbed by staff at several positions around patient being exposed to the ionizing radiation. Previously, the scattered dose absorbed by an infant being administered a chest radiograph was shown to be less than the background radiation (Catherine et al., 2009; Trinh et al., 2010). Similar findings were concluded for the assessment of the scattered dose absorbed by technicians at different distances from X-ray exposure (Chiang et al., 2015).

Nevertheless, the research interest of this study is not on the effectiveness of shielding material in an X-ray room. Rather, the distribution of scattered dose inside the X-ray room is the focus of this study. An X-ray room is known to be commonly equipped with a single X-ray unit. Therefore, the dose distribution inside an X-ray room was not much of a concern previously. Meanwhile, our X-ray room is equipped with a newly-designed Direct Digital Radiography (DDR) system, a Samsung Varian Transbay DDR X-ray machine, model GC85A. This system is equipped with a railing system that moves the X-ray tube 6504 mm in the y-direction and 4010 mm in the x-direction. In accordance with that, the patient exposure can generally be performed at two different bed positions. Even though the single X-ray tube allows a single exposure at a time, simultaneous preparation of another patient for the next exposure will improve the patient throughput and consequently reduce the waiting time. In addition, this imaging room is equipped with two patient couches at a distance of 4000 mm from the X-ray tube position. Accordingly, consideration of the unplanned amount of scattered dose absorbed by patient due to X-ray exposure inside the room is necessary for his/her safety. This is in line with the guidelines set by ICRP on the safe levels of exposure to medical radiation for staff and patients. Therefore, the amount of unnecessary exposure to the patients and the staff should be well observed.

In this study, Monte Carlo modelling of the scattered dose distribution in the multi-bay resuscitation room is performed using MCNP5 code. Wide application of Monte Carlo in medical physics was previously published, which includes dose assessment for many applications related to medical physics (Chiang et al., 2015; Lawrence et al., 2015; McVey & Weatherburn, 2004)**.** The scattered dose from a solid phantom of tissue substitute was calculated as a representative of the scattering behaviour in a patient's body. The calculated data were initially verified against the measurements carried out in the room. The simulation provided new data by calculating the scatter dose distribution inside the imaging room for varying imaging conditions. Nevertheless, the method described in this work is generally applicable for any imaging room.

METHODS

The scattered radiation dose absorbed by patient or staff during a single X-ray examination was calculated inside the multi-bay resuscitation room. Figure 1 shows the schematic illustration of the experimental setup utilising a general-purpose Monte Carlo code version MCNP5. The modelling was designed to be as precise as the real multi-bay resuscitation room in our institution. This precision is necessary to reproduce the same radiation interaction inside the imaging room. This study had modelled 10765 mm x 7868 mm room dimension, patient couches, X-ray beam, and concrete wall. Four patient couches positioned inside the room were modelled via 6 mm thick of water material (Goertz et al., 2015). The distances between the couches were 2990 mm (between T1 and T2) and 3120 mm (between T3 and T4) in the y-direction. Meanwhile, the x-direction distance between the couches were 4000 mm.

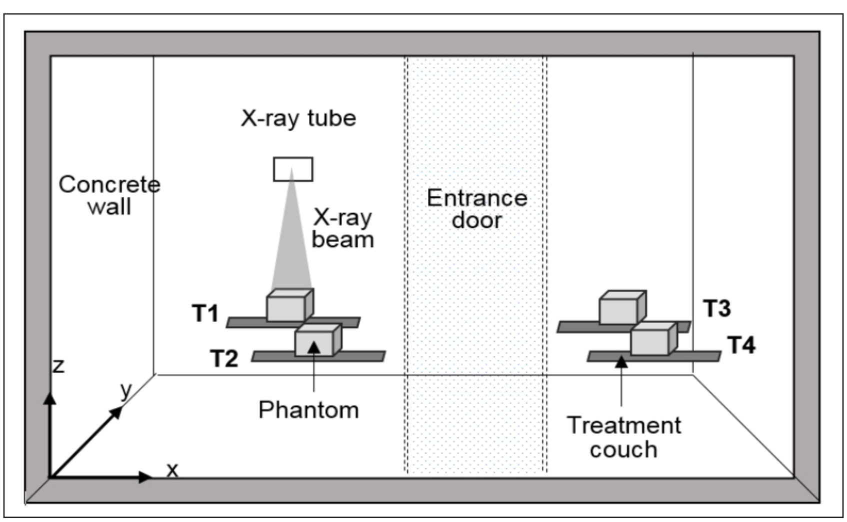


Figure 1. Illustration of the multi-bay resuscitation room modelled using MCNP5 code. T1 and T2 are the couches for X-ray beam exposure, while T3 and T4 are the treatment couches for patients

Three measurement positions were considered during the calculation, labelled as T2, T3, and T4. These positions were selected due to the higher probability that a staff or a patient will be present at the respective positions during imaging. A 30 x 30 x 30 cm solid phantom of tissue substitute was positioned on the patient couch labelled T1, at a source-to-surface distance (SSD) of 100 cm from the X-ray tube focus. An X-ray beam was vertically projected on the upper surface of the phantom. The calculated data were then used to deduce the scattering radiation for the respective clinical settings. The height of the detector was fixed at 90 cm above the floor, which was at the same level of a patient lying on the couch. Hence, measurements were obtained at a scattering angle of 90^{o} with a tube voltage of 125 kV for an incident field area of 900 cm^2. Figure 2 illustrates the experimental setup. The 125 kV energy spectrum was adapted from the Siemens online tool for the simulation of X-ray spectra. The spectrum was collimated into a cone of direction, projected onto the surface of the solid phantom of tissue substitute, modelled to simulate a patient X-ray imaging.

Initially, the calculated data were verified against the measured data obtained using Fluke 481 Radiation Survey meter (David & Otha, 2009; Owusu-Banahene et al., 2018). This step was performed to validate the accuracy of our model as well as the scattered dose estimation. This precision is necessary in order to perform a simulation with similar characteristics as the real room condition. Hence, accurate estimation of the scattering activities in the room could be produced. For each measurement position, the exposure rate at the periphery of the couch was recorded. The survey meter was held at 90 cm height from the floor, to measure the exposure rate at the same level of a patient lying on the couch. The average measured exposure rate was then considered for scattered dose calculation.

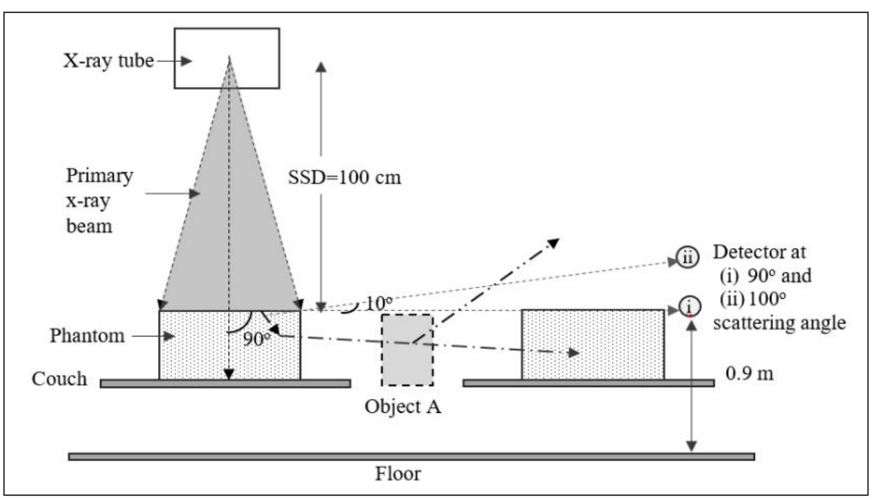


Figure 2. The experimental setup for exposure rate measurement and simulation from a solid phantom of tissue substitute exposed to the X-ray beam. Object A represents the other equipment present in the X-ray room that may lead to other scatter interactions

The mean value of the measured exposure rate was converted to the absorbed dose inside a water phantom. Given, the average energy dissipated in the production of a single ion pair in the air is 34 eV, 1 eV equals to 1.6 x 10^{-19} J, and the charge on a single ion is 1.6 x 10^{-19} C. Therefore, the absorbed dose (in a unit of Gy) is presented by Equation 1, whereby X refers to the exposure in the unit of mR, μ is the mass attenuation coefficient, and 2.58 x 10^{-7} is the mR to C kg^{-1} conversion coefficient (Thomas & Herman, 2009).

$$D_{Exp} = X \times (2.58 \times 10^{-7}) \times 34 \times \left(\frac{\mu_{water}}{\mu_{air}}\right)_{mass} \quad (1)$$

Meanwhile, the dose or energy deposition in the target cell was scored using *F8 and F6 tally in MCNP5 calculation. The *F8 tally gives an average energy deposited in a unit of MeV. The average energy (E) was then converted to absorbed dose by using Equation 2, whereby 1.602 x 10^{-10} is the MeV g^{-1} to Gy conversion factor. The mass (m) of the cell was determined by multiplying the density of material in the cell and the volume of the cell. In addition to that, the total dose in a water phantom was calculated with the consideration of the X-ray tube efficiency, the tube current, and the exposure time. The X-ray tube efficiency for the 125 kVp tube voltage is equal to 0.01 (Hertrich, 2005).

$$D_{MCNP5} = \left[\frac{E}{m} \times (1.602 \times 10^{-10})\right] \quad (2)$$

The F6 tally, which gives data in the unit of MeV/g was converted to the absorbed dose by multiplying with 1.602 x 10^{-10} Gy. Similar consideration of the X-ray tube efficiency, tube current, and exposure time were also calculated. The simulation was performed using 10^6 particles to achieve small relative errors (less than 0.10) and pass the statistical checks performed by the package. The relative error less than 0.10 shows a reliable confidence interval and eventually is a measure of the computational precision. Each of the simulations was then repeated using three different seed numbers to observe the reproducibility of the data. The error (E_r) between the simulated (D_{MCNP5}) and the measured (D_{Exp}) values was calculated through the percentage of deviation between the two values, as described by Equation 3 (Gu´erin & Fakhri, 2008).

$$E_r = \frac{|D_{MCNP5} - D_{Exp}|}{D_{Exp}} \times 100 \quad (3)$$

In this study, we have considered several imaging conditions to be assessed such as the phantom size, the beam collimation size, and the beam projection angles. Figure 3 shows the illustration of the nine beam projection angles modelled using the Monte Carlo code. For the subsequent experiments, the calculations were performed at scattering angles of 90^o and 100^o. The 90^o scattering angle represented the patient lying on the couch, while the

100^o scattering angle represented the scattered dose absorbed by patient or staff standing at their respective positions. Figure 2 illustrates the two scattering angles.

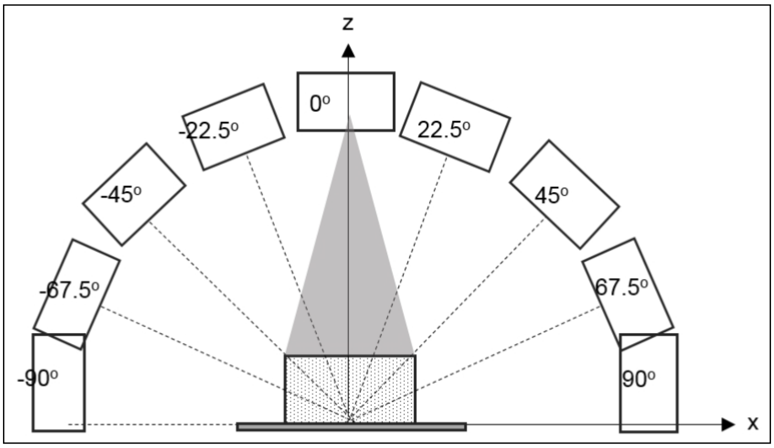


Figure 3. Assuming that the vertical projection was 0^o projection, the source of the beam was angled at 22.5^o step each

In this study, the MCNP5 code was defined to transport the X-ray beam to the phantom volume, which was modelled to represent the patient's body. Therefore, the estimation of the scattered dose that would be absorbed by patient or staff was calculated by the amount of dose absorbed by the solid phantom of tissue substitute. In addition to the photon and phantom interaction that leads to the initiation of the Compton scattering interaction, the presence of other materials inside the room enhances the scatter interaction

Table 1
The density and composition of the materials defined in the MCNP5 code

		Material			
		Air	Water	Concrete	Lead
Density ($g\ cm^{-3}$)		0.001205	1.00	2.30	11.35
Elemental composition (%)	H	-	11.19	1.00	-
	C	0.01	-	0.10	-
	N	75.53	-	-	-
	O	23.18	88.81	52.91	-
	Na	-	-	1.60	-
	Mg	-	-	0.20	-
	Al	-	-	3.39	-
	Si	-	-	33.70	-
	K	-	-	1.30	-
	Ca	-	-	4.40	-
	Fe	-	-	1.40	-
	Ar	1.28	-	-	-
	Pb	-	-	-	100.00

(McVey & Weatherburn, 2004). In our code, only the main equipment necessary for the assessment were considered. Therefore, the difference between the MCNP5 scattered dose and the measured scattered dose was expected. Table 1 tabulates the density and the material composition used in the MCNP5 calculation. The density and the composition of materials used in this simulation were obtained from the National Institute of Standards and Technology (NIST).

RESULTS AND DISCUSSION

The scattered doses calculated from the measured exposure rate are tabulated in Table 2. The scattered doses measured at T3 and T4 were much lower than those measured at T2. Higher scattered dose obtained at T2 was expected, due to the positioning of the X-ray tube which was nearer to the T2 position. According to the inverse square law, the radiation intensity is inversely proportional to the square of the distance from the source. Hence, as T2 was nearer to the X-ray tube, it led the positioned phantom to absorb more radiation and doses compared to T3 and T4. Nevertheless, the value was still lower than the annual dose limit for the public and radiation workers as recommended by the Basic Safety Standard regulations. In addition, this value was lower than the natural background radiation dose (8.493 µGy per day) (Trinh et al., 2010).

Table 2
The absorbed dose calculated from the measured exposure rate at four measurement points around the couch (SD=standard deviation)

	Absorbed dose (µGy)		
	T2	T3	T4
Point 1	5.24	2.81	4.07
Point 2	1.36	2.81	0.87
Point 3	6.21	3.01	2.52
Point 4	9.89	2.81	2.33
Mean	5.68	2.86	2.45
SD	3.51	0.10	1.31

Figure 4 shows the comparison between the scattered doses calculated using *F8 and F6 tally. A small error of 2.22% was observed between the two. Accordingly, subsequent experiments to calculate the scattered doses were performed using *F8 tally only. On the other hand, the data compared between measurements and MCNP5 simulation showed significant difference. Maximum difference of 31.8% was calculated between the two. The overestimation of the MCNP5 absorbed dose was most probably due to the simplicity of the room modelled in MCNP5 code. Although the real room is furnished with equipment such as trolley, chairs, and accessories, our MCNP5 code only considered the main equipment and modality necessary for the assessment. Figure 2 shows the schematic diagram of the interaction between the scattered photon and object A (represents other equipment present inside the room). This illustration is intended to describe the effect of simple geometry modelled in MCNP5 simulation as compared to the actual environment in the X-ray room. Without the object A, the scattered photon may directly hit the phantom next to the phantom exposed to the X-ray beam (shown by dash-dot line). However, with the presence of object A, the photon may scatter and deviate from its trajectory (shown by dash-dot-dot line) and eventually miss the next phantom. Nevertheless, similar trends of dose distribution were observed between the two methods of measurement.

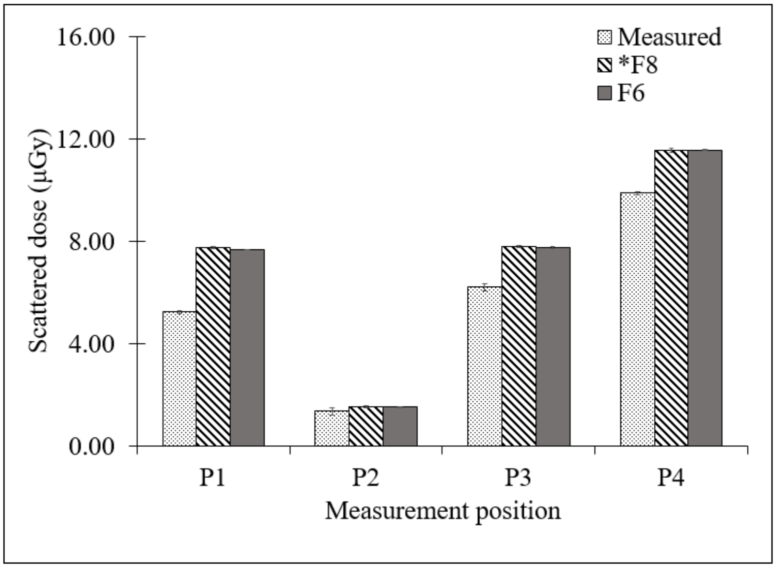


Figure 4. Comparison of the measured and MCNP5 calculated scattered dose. Error bar represents the standard deviation of the data

Although the measurement of exposure rate yielded a low amount of scattered dose inside the imaging room, scattered dose resulted from other exposure settings should be confirmed as well. With respect to that, the MCNP5 simulation was extended to provide calculated scattered dose for other imaging conditions. Figure 5 shows the increment of the scattered dose in correlation with the increment of the phantom size. This increment is expected, considering that the patient's body is known to be the main source of scattering interaction in medical imaging. Referring to the photon interaction cross-section, the probability of Compton interaction is high for the energy range involved in diagnostic radiology imaging. This probability increases as the patient's body size increases, due to the dependency of the Compton interaction on the electron per unit volume (product of physical and electron density). A scattered dose of approximately 21 μGy was calculated at T2 position when a 40 x 40 x 40 cm phantom was exposed to the beam. Indeed, the value was approximately three times higher than the daily background radiation dose. For that reason, greater safety precaution should be considered when imaging larger-sized patients as the body could be a source of intense scatter interaction. On the other hand, calculations revealed that T3 and T4 positions still received lesser amount of scattered dose than the daily background radiation dose, even for the biggest phantom tested.

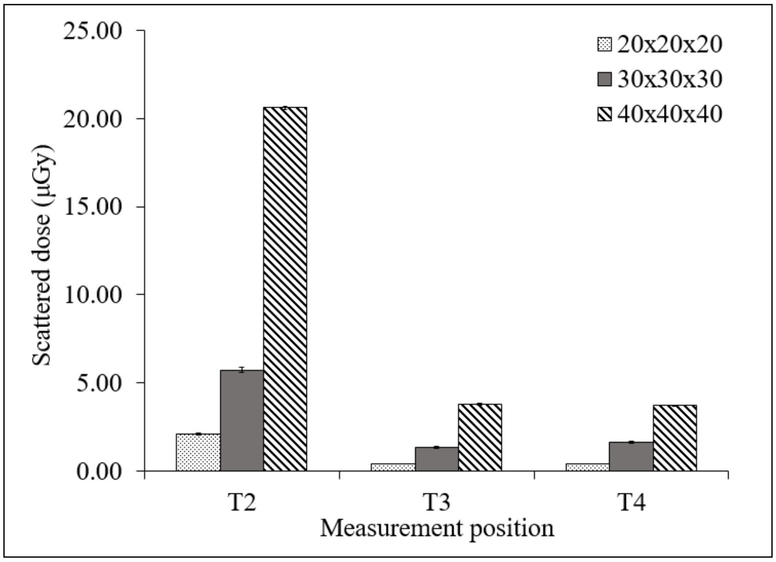


Figure 5. MCNP5 calculated absorbed dose for the different sizes of the phantom. Error bar represents the standard deviation of the data

The comparison shows that the scattered doses calculated at T3 and T4 positions were up to 82% less than the amount calculated at T2 position. In addition, the comparison shows that the scattered doses calculated at T3 and T4 were not significantly varied in the three phantom sizes tested. Nevertheless, the scatter value calculated at T3 was observed to be slightly higher than that calculated at T4. The maximum difference of approximately 16.35% was observed between the two Transbay positions. A similar trend was also observed for the different sizes of beam collimation (Figure 6). Again, T3 recorded higher scattered dose as compared to T4. The maximum difference of 25.62% was recorded between the two. For the 30 x 30 cm beam collimation, greater attention should be given to the T2 position considering that a single exposure could result in a scattered dose that exceeds the daily background radiation.

Calculation of the scattered dose was then extended to the different angles of the X-ray tube. The source angled at 90^{o} in the x and y-directions was considered to evaluate the scattered dose to the three bed positions. These angles were considered to simulate the different beam projections that could be performed during imaging. Figure 7 shows the variation of the scattered dose absorbed by a patient at a 90^{o} scattering angle for the

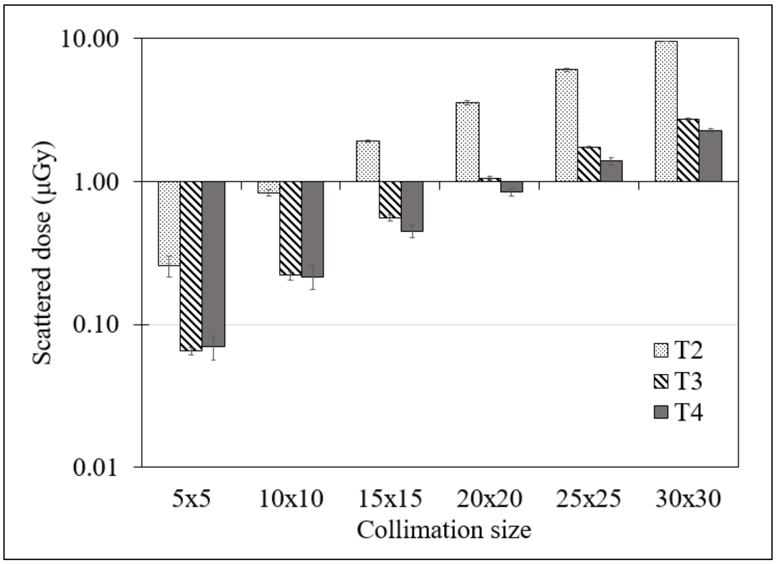


Figure 6. MCNP5 calculated scattered dose for the beam collimation in the range of 5x5 to 30x30. Error bar represents the standard deviation of the data

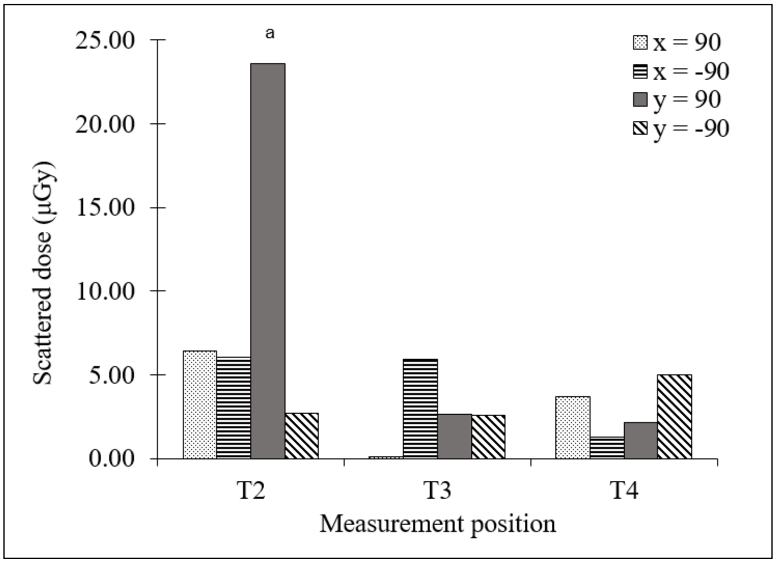
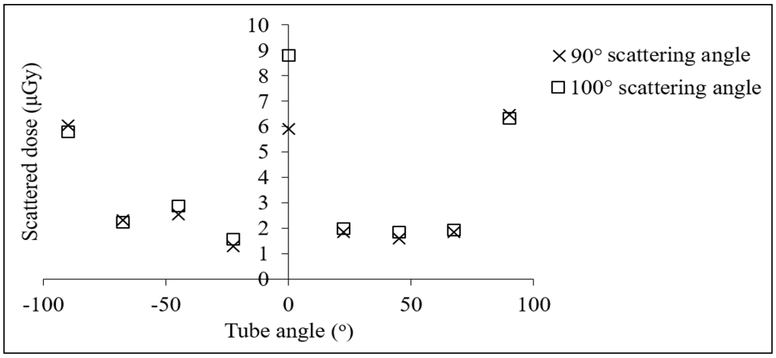
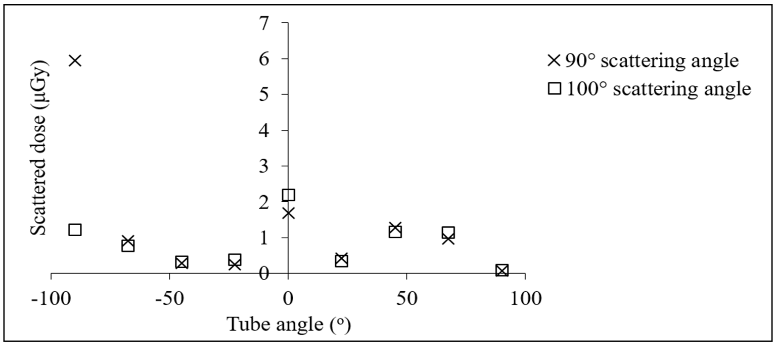


Figure 7. The scattered dose obtained from a collimated cone source angled at 90^o in the x and y-directions

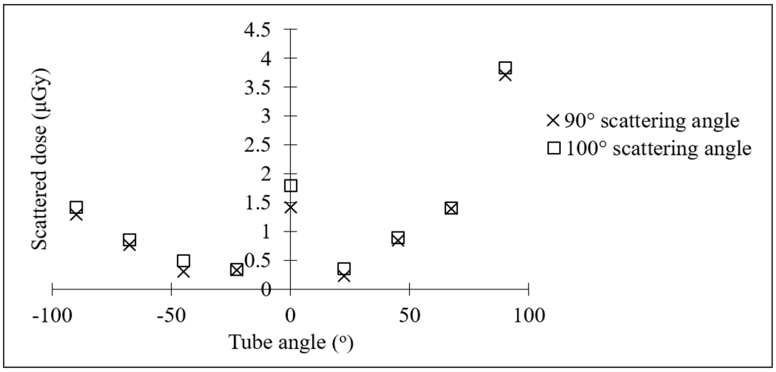
investigated angles. The result shows that a relatively low scattered dose was measured at the three positions. The maximum scattered dose of approximately 6 µGy was measured for the tested angles. Nevertheless, the projection parallel to T2 position could result in significant amount of scattered dose to the patient unplanned for exposure at T2 position (denoted by [a] in Figure 7). However, it should be noted that this projection has never been implemented in our institution. In conclusion, the amount of scattered dose measured for the commonly practiced exposure angles tested is safe and even lower than the annual dose limit for the public and radiation workers. The value is also lower than the natural background radiation dose. To evaluate the scattered doses deposited within the phantom's volume at two patient positioning, 90° and 100° scattering angles were assessed. As shown in Figure 8(a) to Figure 8(c), higher amount of scattered doses were calculated when the projection



(a)



(b)



(c)

Figure 8. The scattered dose calculated for the range of 0° to 180° tube angles in the x-direction, while the measurements were calculated at 90° and 100° scattering angles at (a) T2 (b) T3 and (c) T4

was made at 0^o and 90^o (both positive and negative directions) tube angles. The maximum difference of 48.65% was calculated between the two. Meanwhile, a small difference was observed between the two scattering angles for the tube angles of 22.5^o to 67.5^o (in either positive or negative direction). The high value of the scattered dose was calculated at T3 when the exposure was performed at -90^o tube angle due to the direct projection of the beam towards the T3 position. However, this projection is also never being practiced in our institution. Therefore, the data showed that the amounts of the scattered doses were within safe level for all angles assessed. For the 100^o scattering angle, the amount of scattered dose at T2 may slightly exceed the daily background radiation dose when the exposure was performed at 0^o tube angle.

With the concern that T2 had consistently received the highest scattered dose compared to T3 and T4 positions, a lead shielding was proposed to be positioned between the two to reduce such exposure. The portable lead shield with the distance range of 20 cm to 100 cm to the X-ray tube position was modelled using MCNP5 code. The calculation revealed that no significant difference was observed among the tested lead shield positions. Nevertheless, the positioning of the lead shield between the X-ray tube and T2 was able to significantly reduce the scattered dose, up to 98%. Accordingly, the positioning of a lead shield is highly recommended, especially for exposure settings that lead to higher scatter values.

CONCLUSION

In this study, characterisation of the scatter dose distribution in the multi-bay resuscitation room has been performed. This characterisation is important due to the placement of multiple patient couches in the respective room. Hence, the amount of unnecessary, unintended exposure absorbed by patient and staff during the imaging procedure should be well estimated. The investigation was conducted using MCNP5 code. The assessment revealed that the highest scattered dose was calculated at T2. The main reason for this finding was due to the shorter distance between X-ray tube and T2. Yet, the calculated value was still lower than the dose limit for the public and radiation workers as recommended by the Basic Safety Standard regulations. The value was also lower than the daily background radiation dose of 8.22 μGy reported by NCRP Report No. 160. Nevertheless, it is strongly recommended that appropriate lead shielding should be considered when performing the imaging procedure at T1 while another patient is positioned at T2. In addition, although T3 and T4 received relatively small amounts of the scattered doses (much lower than the daily background radiation), the placement of protective lead shield should be considered. This is because the calculation only considered a single exposure setting. In a busy hospital, the number of imaging procedures performed daily is high. Hence, the cumulative dose to paramedic staff and patients may be very high.

ACKNOWLEDGEMENTS

This study was supported by the USM Short Term Grant (304/PPSK/6315167). The authors would like to thank the School of Health Sciences, USM and the Director of the Hospital Universiti Sains Malaysia (USM), Kubang Kerian, Kelantan for granting the permission to the investigators to use space and assets belonging to the hospital during the process of conducting the research. Special thanks goes to the staff of Hospital USM who had relentlessly assisted in making the research work successful.

REFERENCES

Archer, B. R., Fewell, T. R., Conway, B. J., & Quinn, P. W. (1994). Attenuation properties of diagnostic X-ray shielding materials. *Medical Physics, 21*, 1499-1507.

Catherine, M. H., John, H. M., Homer, C. T., & John, B. K. (2009). Radiation dose from initial trauma assessment and resuscitation: review of the literature. *Canadian Journal of Surgery, 52*(2), 147-152.

Chiang, H. W., Liu, Y. L., Chen, T. R., Chen, C. L., Chiang, H. J., & Chao, S. Y. (2015). Scattered radiation doses absorbed by technicians at different distances from X-ray exposure: Experiments on prosthesis. *Bio-Medical Materials and Engineering, 26*, S1641-S1650.

David, A. S., & Otha, W. L. (2009). National council on radiation protection and measurements report shows substantial medical exposure increase. *Radiology, 253*, 293-296.

Gu´erin, B., & Fakhri, G. E. (2008). Realistic PET Monte Carlo simulation with pixelated block detectors, light sharing, random coincidences and dead-time modeling. *IEEE Transactions on Nuclear Science, 55*(3), 942-952.

Goertz, L., Tsiamas, P., Karellas, A., Sajo, E., & Zygmanski, P. (2015). Monte Carlo simulation of a prototypical patient dosimetry system for fluoroscopic procedures. *Physics in Medicine and Biology, 60*(15), 5891-5909.

Hertrich, P. H. (2005). *Practical radiography: Principles and applications*. Erlangen, Germany: Publicis Corporate Publishing.

Lawrence, W. C. C., Chan, T. P., Ben, T. F. C., Kelvin, M., & Karl, K. L. F. (2015). Simulation, visualization and dosimetric validation of scatter radiation distribution under fluoroscopy settings. *Journal of Biomedical Engineering and Informatics, 1*(1), 93-102.

McVey, G., & Weatherburn, H. (2004). A study of scatter in diagnostic X-ray rooms. *The British Journal of Radiology, 77*, 28-38.

Owusu-Banahene, J., Darko, E. O., Charles, D. F., Maruf, A., Hanan, I., & Amoako, G. (2018). Scatter radiation dose assessment in the radiology department of cape coast teaching hospital-Ghana. *Open Journal of Radiology, 8*(4), 299-306.

Simpkin, D. J. (1989). Shielding requirements for constant potential diagnostic X-ray beams determined by a Monte Carlo calculation. *Health Physics, 56*, 151-164.

Simpkin, D. J. (1995). Transmission data for shielding diagnostic X-ray facilities. *Health Physics, 68*, 704-709.

Thomas, E. J., & Herman, C. (2009). *Introduction to health physics*. New York, NY: McGraw-Hill Medical.

Trinh, A. M., Schoenfeld, A. H., & Levin, T. L. (2010). Scatter radiation from chest radiographs: is there a risk to infants in a typical NICU? *Pediatric Radiology, 40*, 704-707.

REFEREES FOR THE PERTANIKA
JOURNAL OF SCIENCE AND TECHNOLOGY

VoL. 27 (4) Oct. 2019

The Editorial Board of the Journal of Science and Technology wishes to thank the following:

Abdul Aziz Hairuddin
(UPM, Malaysia)

Abdullah Hisam Omar
(UTM, Malaysia)

Adi Ainurzaman Jamaludin
(UM, Malaysia)

Adibah Mohd Amin
(UPM, Malaysia)

Ahmad Baharuddin Abdullah
(USM, Malaysia)

Ahmad Farid Abu Bakar
(UM, Malaysia)

Ahmad Izani Md Ismail
(USM, Malaysia)

Ahmad Taufek Abdul Rahman
(UiTM, Malaysia)

Ahmad Zahran Md Khudzari
(UTM, Malaysia)

Aishah Hani Azil
(PPUKM, Malaysia)

Alfian Abdul Halin
(UPM, Malaysia)

Ali Hassan Mohamed Murid
(UTM, Malaysia)

Amin Ismail
(UPM, Malaysia)

Amir Izzwan Zamri
(UMT, Malaysia)

Amit Bairagi
(Patankar Hospital, India)

Anusha Achuthan
(USM, Malaysia)

Arien Heryansyah
(UTM, Malaysia)

Arifah Bahar
(UTM, Malaysia)

Aweng Eh Rak
(UMK, Malaysia)

Azeana Zahari
(UM, Malaysia)

Azhar Abdul Halim
(UKM, Malaysia)

Azhari
(Universiti Malikussaleh, Indonesia)

Azizan As'arry
(UPM, Malaysia)

Azizul Azhar Ramli
(UTHM, Malaysia)

Azizul Helmi Sofian
(UMP, Malaysia)

Azmah Hanim Mohamed Ariff
(UPM, Malaysia)

Azni Zain Ahmed
(UiTM, Malaysia)

Chai Chang Saar
(UoRM, Malaysia)

Che Badariah Ab Aziz
(USM, Malaysia)

Che Hassan Che Haron
(UKM, Malaysia)

Che Maznah Mat Isa
(UiTM, Malaysia)

Che Zulzikrami Azner Abidin
(UniMAP, Malaysia)

Cheng Chin Kui
(UMP, Malaysia)

Chonticha Kaewanuchit
(ARU, Thailand)

Chua Han Bing
(Curtin University, Malaysia)

Chuah Cheng Hock
(UM, Malaysia)

Darius El Pebrian
(UiTM, Malaysia)

Desa Ahmad
(UPM, Malaysia)

Dillip Kumar Ghose
(NIT, India)

Duratul Ain Tholibun
(UITM, Malaysia)

Dzati Athiar Ramli
(USM, Malaysia)

Dzun Noraini Jimat
(IIUM, Malaysia)

Eris Elianddy Supeni
(UPM, Malaysia)

Ezamin Abdul Rahim
(UPM, Malaysia)

Fahmiruddin Esa
(UTHM, Malaysia)

Fahrul Hakim Ayob
(UPM, Malaysia)

Fairolniza Mohd Shariff
(UPM, Malaysia)

Faradina Merican Mohd Sidik Merican
(USM, Malaysia)

Fatimah Khalid
(UPM, Malaysia)

Florentin Smarandache
(UNM, USA)

Gang Li
(Deakin University, Australia)

Geetha Annavi
(UPM, Malaysia)

Ghufran Redzwan
(UM, Malaysia)

Hafiz Rashidi Harun @ Ramli
(UPM, Malaysia)

Hashibah Hamid
(UUM, Malaysia)

Hiewa Dyary
(University of Sulaimani, Iraq)

Hizlinda Tohid
(UKM, Malaysia)

Irfan Ahmed Shaikh
(SAU, Pakistan)

Isa Bala Muhammad
(FUTMinna, Nigeria)

Iskandar Ishak
(UPM, Malaysia)

Jasronita Jasni
(UPM, Malaysia)

Jolius Gimbun
(UMP, Malaysia)

Juliana Aida Abu Bakar
(UUM, Malaysia)

Khairil Anas Md Rezali
(UPM, Malaysia)

Khaironizam Md Zin
(USM, Malaysia)

Khairul Hamimah Abas
(UTM, Malaysia)

Law Ming Chiat
(Curtin University, Malaysia)

Lee Sau Har
(Taylor's University, Malaysia)

Lili Nurliyana Abdullah
(UPM, Malaysia)

Lim Chin Wai
(UNITEN, Malaysia)

Lim Soh Fong
(UNIMAS, Malaysia)

Liyaqat Bashir
(LPU, India)

Low Siew Chun
(USM, Malaysia)

Maizatul Akmar Ismail
(UM, Malaysia)

Malina Osman
(UPM, Malaysia)

Marcus Jopony
(UMS, Malaysia)

Mariam Adawiah Dzulkifli
(IIUM, Malaysia)

Marsyita Hanafi
(UPM, Malaysia)

Masnida Hussin
(UPM, Malaysia)

Md Azlin Md Said
(USM, Malaysia)

Mohamed Abd Rahman
(IIUM, Malaysia)

Mohammad Abdul Mojid
(BAU, Malaysia)

Mohammed Alias Yusof
(UPNM, Malaysia)

Mohd Azlan Mohd Ishak
(UiTM, Malaysia)

Mohd Azri Mohd Izhar
(UTM, Malaysia)

Mohd Dilshad Ansari
(CMRCET, India)

Mohd Fadzil Mohd Akhir
(UMT, Malaysia)

Mohd Nashrul Mohd Zubir
(UM, Malaysia)

Mohd Nazim Mohtar
(UPM, Malaysia)

Mohd Sapuan Salit
(UPM, Malaysia)

Mohd Shahrul Nizam Mohd Danuri
(KUIS, Malaysia)

Mohd Shamsul Anuar
(UPM, Malaysia)

Mohd Shareduwan Mohd Kasihmuddin
(USM, Malaysia)

Mohd Sobri Idris
(UNIMAP, Malaysia)

Mohd Yusoff Ishak
(UPM, Malaysia)

Mohd Zulkhairi Mohd Yusoff
(UPM, Malaysia)

Muhamad Shakirin Mispan
(UM, Malaysia)

Muhammad Nur Salihin Yusoff
(USM, Malaysia)

Muhammad Syafiq Hazwan Ruslan
(UTP, Malaysia)

Muskhazli Mustafa
(UPM, Malaysia)

Muzzneena Ahmad Mustapha
(UKM, Malaysia)

Nafarizal Nayan
(UTHM, Malaysia)

Najla Mohamed Abushaala
(University of Tripoli, Libya)

Nazmi Mat Nawi
(UPM, Malaysia)

Ngah Zasmy Unyah
(UPM, Malaysia)

Nik Norsyahariati Nik Daud
(UPM, Malaysia)

Noorlaila Ahmad
(UiTM, Malaysia)

Nor Ashidi Mat Isa
(USM, Malaysia)

Nor Azuana Ramli
(UniKL, Malaysia)

Nor Azura Husin
(UPM, Malaysia)

Nor Hayati Othman
(USM, Malaysia)

Nora Muda
(UKM, Malaysia)

Norashiqin Misni
(UPM, Malaysia)

Noridayu Manshor
(UPM, Malaysia)

Normala Halimoon
(UPM, Malaysia)

Norshita Mat Nayan
(UKM, Malaysia)

Nur Azida Che Lah
(UniKL, Malaysia)

Nur Shaylinda Mohd Zin
(UTHM, Malaysia)

Nurul Hashimah Ahamed Hassain Malim
(USM, Malaysia)

Othman A. Karim
(UKM, Malaysia)

Perumal Kumar
(Curtin University, Malaysia)

Poh Mau Ern
(UM, Malaysia)

Rabiah Abdul Kadir
(UKM, Malaysia)

Rafidah Hanim Shueb
(USM, Malaysia)

Rasha Thabit Mohammed
(ARUC, Iraq)

Rashima Mahajan
(MRIIRS, India)

Riza Wirawan
(UNJ, Indonesia)

Rohazila Mohamad Hanafiah
(USIM, Malaysia)

Rosiah Rohani
(UKM, Malaysia)

Rosilah Hassan
(UKM, Malaysia)

Roslaini Abd Majid
(UPM, Malaysia)

Rozi Abdullah
(USM, Malaysia)

Ruslan Rainis
(USM, Malaysia)

S.M. Anisuzzaman
(UMS, Malaysia)

Safaa A Khadum
(Qadisiyah University, Iraq)

Saleha Shamsudin
(UNIMAP, Malaysia)

Salina Abdul Samad
(UKM, Malaysia)

Sangwan Kuldip Singh
(BITS, India)

Satish Kumar Singh
(JUET, India)

Sayed Inayatullah Shah
(IIUM, Malaysia)

Sharifah Akmam Syed Zakaria
(USM, Malaysia)

Shengwei Wang
(PolyU, Hong Kong)

Siti Nurbaya Oslan
(UPM, Malaysia)

Sri Harini
(UIN MMI, Indonesia)

Suhaili Abu Bakar @ Jamaludin
(UPM, Malaysia)

Syamsul Rizal Abd Shukor
(USM, Malaysia)

Syed Abdul Rahman Al-Haddad Syed Mohamed
(UPM, Malaysia)

Tanuj Kumar
(LPU, India)

Tay Chia Chay
(UiTM, Malaysia)

Thaned Rojsiraphisal
(CMU, Thailand)

Tye Ching Thian
(USM, Malaysia)

Waida Ismail
(USIM, Malaysia)

Wan Azizun Wan Adnan
(UPM, Malaysia)

Wan Nurul Huda Wan Zainal
(UMP, Malaysia)

Waqar Asrar
(IIUM, Malaysia)

Yap Wei Boon
(UKM, Malaysia)

Yasmin Mustapha Kamil
(UPM, Malaysia)

Yeoh Wei Ming
(UTAR, Malaysia)

Yoon Li Wan
(Taylor's University, Malaysia)

Zahiruddin Othman
(USM, Malaysia)

Zainul Amiruddin Zakaria
(UPM, Malaysia)

Zamzuri Hamedon
(UMP, Malaysia)

Zarani Mat Taher
(UTM, Malaysia)

Zarul Hazrin Hashim
(USM, Malaysia)

Zubaidah Ismail
(UM, Malaysia)

Zulkarnain Hassan
(UNIMAP, Malaysia)

Zurina Zainal Abidin
(UPM, Malaysia)

ARU	*– Ayutthaya Rajabhat University*
ARUC	*– Al Rasheed University College*
BITS	*– Birla Institute of Technology and Science*
CMRCET	*– CMR College of Engineering and Technology*
CMU	*– Chiangmai University*
FUTMinna	*– Federal University of Technology Minna*
IIUM	*– International Islamic University Malaysia*
JUET	*– Jaypee University of Engineering and Technology*
KUIS	*– Kolej Universiti Islam Antarabangsa Selangor*
LPU	*– Lovely Professional University*
MRIIRS	*– Manav Rachna International Institute Research & Studies*
NIT	*– National Institutes of Technology*
PolyU	*– The Hong Kong Polytechnic University*
PPUKM	*– Pusat Perubatan Universiti Kebangsaan Malaysia*
SAU	*– Sindh Agriculture University*
UIN MMI	*– Universitas Islam Negeri Maulana Malik Ibrahim*
UiTM	*– Universiti Teknologi MARA*
UKM	*– Universiti Kebangsaan Malaysia*
UM	*– Universiti Malaya*
UMP	*– Universiti Malaysia Pahang*
UMS	*– Universiti Malaysia Sabah*
UMS.ID	*– Universitas Muhammadiyah Surakarta, Indonesia*
UMT	*– Universiti Malaysia Terengganu*
UniKL	*– University Kuala Lumpur*
UNIMAP	*– Universiti Malaysia Perlis*
UNIMAS	*– Universiti Malaysia Sarawak*
UNITEN	*– Universiti Tenaga Nasional*
UNJ	*– UNIVERSITAS NEGERI JAKARTA*
UNM	*– University of New Mexico*
UoRM	*– University of Reading Malaysia*
UPM	*– Universiti Putra Malaysia*
UPNM	*– Universiti Pertahanan Nasional Malaysia*
USIM	*– Universiti Sains Islam Malaysia*
USM	*– Universiti Sains Malaysia*
UTAR	*– Universiti Tunku Abdul Rahman*
UTHM	*– Universiti Tun Hussein Onn Malaysia*
UTM	*– Universiti Teknologi Malaysia*
UTP	*– Universiti Teknologi Petronas*
UUM	*– Universiti Utara Malaysia*

While every effort has been made to include a complete list of referees for the period stated above, however if any name(s) have been omitted unintentionally or spelt incorrectly, please notify the Chief Executive Editor, *Pertanika* Journals at executive_editor.pertanika@upm.edu.my

Pertanika Journals

Our goal is to bring high quality research to the widest possible audience

INSTRUCTIONS TO AUTHORS

(Manuscript Preparation & Submission Guide)

Revised: June 2019

Please read the Pertanika guidelines and follow these instructions carefully. The Chief Executive Editor reserves the right to return manuscripts that are not prepared in accordance with these guidelines.

MANUSCRIPT PREPARATION

Manuscript Types

Pertanika accepts submission of mainly **four** types of manuscripts for peer-review.

1. Regular article

Regular articles are full-length original empirical investigations, consisting of introduction, materials and methods, results and discussion, conclusions. Original work must provide an explanation on research results that contain new and significant findings. Analysis and Discussion must be supported with relevant references.

Size: Generally, these are expected to be between **6 and 12 journal pages or not exceeding 6000 words** (excluding the abstract, references, tables and/or figures), a maximum of **80 references**, and **an abstract of 100–200 words.**

2. Review article

These report critical evaluation of materials about current research that has already been published by organizing, integrating, and evaluating previously published materials. It summarizes the status of knowledge and outline future directions of research within the journal scope. Review articles should aim to provide systemic overviews, evaluations and interpretations of research in a given field. Re-analyses as meta-analysis and systemic reviews are encouraged. The manuscript title must start with "Review Article:".

Size: Generally, these are expected to be between **8 and 12 journal pages or not exceeding 6000 words** (excluding the abstract, references, tables and/or figures), a maximum of **80 references,** and **an abstract of 100–200 words.**

3. Short communications

They are timely and brief. These are suitable for the publication of significant technical advances and may be used to:

(a) report new developments, significant advances and novel aspects of experimental and theoretical methods and techniques which are relevant for scientific investigations within the journal scope;

(b) report/discuss on significant matters of policy and perspective related to the science of the journal, including 'personal' commentary;

(c) disseminate information and data on topical events of significant scientific and/or social interest within the scope of the journal.

Size: These are usually between **2 and 4 journal pages** and have a maximum of **3 figures and/or tables, from 8 to 20 references, and an abstract length not exceeding 100 words.** Information must be in short but complete form and it is not intended to publish preliminary results or to be a reduced version of Regular or Rapid Papers.

4. Others

Brief reports, case studies, comments, concept papers, letters to the editor, and replies on previously published articles may be considered.

Language Accuracy

Pertanika **emphasizes** on the linguistic accuracy of every manuscript published. Articles must be in **English** and they must be competently written and argued in clear and concise grammatical English. Contributors are strongly advised to have the manuscript checked by a colleague with ample experience in writing English manuscripts or a competent English language editor.

Author(s) **may be required to provide a certificate** confirming that their manuscripts have been adequately edited. **All editing costs must be borne by the author(s)**.

Linguistically hopeless manuscripts will be rejected straightaway (e.g., when the language is so poor that one cannot be sure of what the authors really mean). This process, taken by authors before submission, will greatly facilitate reviewing, and thus publication if the content is acceptable.

MANUSCRIPT FORMAT

The paper should be submitted in one column format with at least 4cm margins and 1.5 line spacing throughout. Authors are advised to use Times New Roman 12-point font and *MS Word* format.

1. *Manuscript Structure*

Manuscripts in general should be organised in the following order:

- **Page 1:** ***Running title***

This page should **only** contain the running title of your paper. The running title is an abbreviated title used as the running head on every page of the manuscript. The running title **should not exceed 60 characters, counting letters and spaces.**

- **Page 2:** ***Author(s) and Corresponding author information***

General information: This page should contain the **full title** of your paper **not exceeding 25 words**, with name(s) of all the authors, institutions and corresponding author's name, institution and full address (Street address, telephone number (including extension), hand phone number, and e-mail address) for editorial correspondence. **First and corresponding authors must be clearly indicated.**

Authors' name: The names of the authors should be named **in full without academic titles.** For Asian (Chinese, Korean, Japanese, Vietnamese), please write first name and middle name before surname (family name). The last name in the sequence is considered the surname.

Authors' addresses: Multiple authors with different addresses must indicate their respective addresses separately by superscript numbers.

Tables / figures list: A **list** of number of **black and white / colour figures and tables** should also be indicated on this page. See ***"6. Figures & Photographs"*** for details.

Example (page 2):

In vivo Fecundity Evaluation of Phaleria macrocarpa Extract Supplementation in Male Adult Rats

Sui Sien Leong[1] and Mohamad Aziz Dollah[2]*

[1]Department of Animal Sciences and Fishery, Universiti Putra Malaysia, 97008 Bintulu, Sarawak, Malaysia
[2]Department of Biomedical Sciences, Universiti Putra Malaysia, 43400 Serdang, Malaysia

leongsuisien@upm.edu.my (Sui Sien Leong), Contact number
azizdollah@gmail.com (Mohamad Aziz Dollah), Contact number
**Corresponding author*

List of Table / Figure:
Table 1.
Figure 1.

- **Page 3: *Abstract***

This page should **repeat** the **full title** of your paper with only the **Abstract** (the abstract should be less than 250 words for a Regular Paper and up to 100 words for a Short Communication), **and Keywords.**

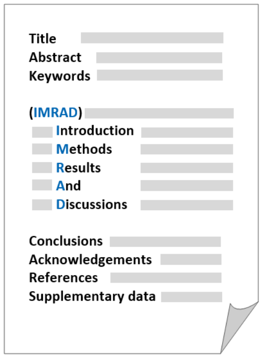
Keywords: Not more than 8 keywords in alphabetical order must be provided to describe the contents of the manuscript.

- **Page 4: *Introduction***

This page should begin with the **Introduction** of your article and followed by the rest of your paper.

2. *Text*

Regular Papers should be prepared with the headings *Introduction, Materials and Methods, Results and Discussion*, *Conclusions, Acknowledgements, References, and Supplementary data* (if any) in this order. The literature review may be part of or separated from Introduction.



MAKE YOUR ARTICLES AS CONCISE AS POSSIBLE

Most scientific papers are prepared according to a format called IMRAD. The term represents the first letters of the words Introduction, Materials and Methods, Results, And, Discussion. It indicates a pattern or format rather than a complete list of headings or components of research papers; the missing parts of a paper are: Title, Authors, Keywords, Abstract, Conclusions, and References. Additionally, some papers include Acknowledgments and Appendices.

The Introduction explains the scope and objective of the study in the light of current knowledge on the subject; the Materials and Methods describes how the study was conducted; the Results section reports what was found in the study; and the Discussion section explains meaning and significance of the results and provides suggestions for future directions of research. The manuscript must be prepared according to the Journal's instructions to authors.

3. *Levels of Heading*

Level of heading	Format
1st	**LEFT, BOLD, UPPERCASE**
2nd	**Flush left, Bold, Capitalise each word**
3rd	**Bold, Capitalise each word, ending with .**
4th	***Bold italic*, Capitalise each word, ending with .**

4. *Equations and Formulae*

These must be set up clearly and should be typed double spaced. Numbers identifying equations should be in square brackets and placed on the right margin of the text.

5. *Tables*

- All tables should be prepared in a form consistent with recent issues of Pertanika and should be numbered consecutively with Roman numerals (Table 1, Table 2).
- A brief title should be provided, which should be shown at the top of each table **(APA format):**

Example: Table 1

PVY infected Nicotiana tabacum *plants optical density in ELISA*

- Explanatory material should be given in the table legends and footnotes.
- Each table should be prepared on a new page, embedded in the manuscript.
- Authors are advised to keep backup files of all tables.

***** Please submit all tables in Microsoft word format only - because tables submitted as image data cannot be edited for publication and are usually in low-resolution.***

6. *Figures & Photographs*

- Submit an **original** figure or photograph.
- Line drawings must be clear, with high black and white contrast.
- Each figure or photograph should be prepared on a new page, embedded in the manuscript for reviewing to keep the file of the manuscript under 5 MB.
- These should be numbered consecutively with Roman numerals (Figure 1, Figure 2).
- Provide a brief title, which should be shown at the bottom of each table **(APA format):**

  Example: Figure 1. PVY-infected *in vitro* callus of *Nicotiana tabacum*

- If a Figure has been previously published, acknowledge the original source and submit written permission form the copyright holder to reproduce the material.
- Authors are advised to keep backup files of all figures.

***** Figures or photographs must also be submitted separately as TIFF, JPEG, because figures or photographs submitted in low-resolution embedded in the manuscript cannot be accepted for publication. For electronic figures, create your figures using applications that are capable of preparing high resolution TIFF files.***

7. *References*

References begin on their own page and are listed in alphabetical order by the first author's last name. Only references cited within the text should be included. All references should be in 12-point font and double-spaced.

NOTE: When formatting your references, please follow the ***Pertanika***-**APA-reference style** (6th Edition) *(refer to examples)*. Ensure that the references are strictly in the journal's prescribed style, failing which your article will **not be accepted for peer-review**. You may refer to the *Publication Manual of the American Psychological Association* for further details https://apastyle.apa.org/.

Examples of reference style are given below:

Books

Books	**Insertion in Text**	**In Reference List**
Book with 1-2 authors	**Information prominent' (the author's name is within parentheses):** … (Cochrane, 2007) … Or **'Author prominent' (the author's name is outside the parentheses):** Cochrane (2007) …	Cochrane, A. (2007). *Understanding urban policy: A critical approach*. Malden, MA: Blackwell Publishing.
Book with 3 or more authors (Pertanika's format)	*For all in-text references, list only the firrst author's family name and followed by* 'et al.' **Information prominent' (the author's name is within parentheses):** … (Seeley et al., 2011) … Or **'Author prominent' (the author's name is outside the parentheses):** Seeley et al. (2011) …	Seeley, R., VanPutte, C., Regan, J., & Russo, A. (2011). *Seeley's anatomy & physiology*. New York, NY: McGraw-Hill.

Books	Insertion in Text	In Reference List
Book with 6-7 authors	*For all in-text references, list only the firrst author's family name and followed by* 'et al.' **Information prominent' (the author's name is within parentheses):** … (Bulliet et al., 2011) … Or **'Author prominent' (the author's name is outside the parentheses):** Bulliet et al. (2011) …	Bulliet, R. W., Crossley, P. K., Headrick, D. R., Hirsch, S. W., Johnson, L. L., & Northrup, D. (2011). *The earth and its peoples: A global history* (5th ed.). Boston, MA: Wadsworth.
Book with more than 8 authors	*For all in-text references, list only the first author's family name and followed by* 'et al.' A recent study (Edge et al., 2011) concluded that… Or Edge et al. (2011) concluded that **….**	For books with eight or more authors, please follow the guidelines for journal articles with eight or more authors.
Chapter in edited book	**Information prominent' (the author's name is within parentheses):** … (Richards, 1997) … Or **'Author prominent' (the author's name is outside the parentheses):** Richards (1997) …	Richards, K. C. (1997). Views on globalization. In H. L. Vivaldi (Ed.), *Australia in a global world* (pp. 29-43). Sydney, Australia: Century.
e-book/online book	**Information prominent' (the author's name is within parentheses):** … (Niemann et al., 2004) … Or **'Author prominent' (the author's name is outside the parentheses):** Niemann (2004) …	Niemann, S., Greenstein, D., & David, D. (2004). *Helping children who are deaf: Family and community support for children who do not hear well.* Retrieved June 1, 2019, from http://www.hesperian.org/ publications_download_deaf.php Schiraldi, G. R. (2001). *The post-traumatic stess disorder sourcebook: A guide to healing, recovery, and growth* [Adobe Digital Editions version]. doi:10.1036/0071393722
Editor	**Information prominent' (the author's name is within parentheses):** … (Zairi, 1999) … Or **'Author prominent' (the author's name is outside the parentheses):** Zairi (1999) …	Zairi, M. (Ed.). (1999). *Best practice: Process innovation management.* Oxford, United Kingdom: Butterworth- Heinemann.

Books	Insertion in Text	In Reference List
Several works by same author in the same year	**Information prominent' (the author's name is within parentheses):** … (Fullan, 1996a, 1997b) … Or **'Author prominent' (the author's name is outside the parentheses):** Fullan (1996a, 1996b) …	Fullan, M. (1996a). Leadership for change. In *International handbook for educational leadership and administration*. New York, NY: Kluwer Academic . Fullan, M. (1996b). *The new meaning of educational change.* London, United Kingdom: Casell.
Several authors, different years referred to collectively in your work	**List sources alphabetically by family name in the in-text reference in the order in which they appear in the Reference List.** The cyclical process (Carr & Kemmis, 1986; Dick, 2000) suggests…	Carr, W., & Kemmis, S. (1986). *Becoming critical: Education knowledge and action research.* London, United Kingdom: Falmer Press. Dick, B. (2000). *A beginner's guide to action research.* Retrieved June 1, 2019, from http://www.scu.edu.au/schools/gcm/ar/arp/guide.html

Journals

Journals	Insertion in Text	In Reference List
Journal article with 1-2 authors	**Information prominent' (the author's name is within parentheses):** … (Kramer & Bloggs, 2002)… Or **'Author prominent' (the author's name is outside the parentheses):** Kramer and Bloggs (2002) …	Kramer, E., & Bloggs, T. (2002). On quality in art and art therapy. *American Journal of Art Therapy*, 40, 218-231.
Journal article with 3 or more authors (Pertanika's format)	*For all in-text references, list only the firrst author's family name and followed by* 'et al.' **Information prominent' (the author's name is within parentheses):** … (Erlo et al., 2008) … Or **'Author prominent' (the author's name is outside the parentheses):** Erlo et al. (2008) …	Elo, A., Ervasti, J., Kuosma, E., & Mattila, P. (2008). Evaluation of an organizational stress management program in a municipal public works organization. *Journal of Occupational Health Psychology*, 13(1), 10-23. doi: 10.1037/1076-8998.13.1.10

Journal article with 6 - 7 authors	*For all in-text references, list only the firrst author's family name and followed by* 'et al.' **Information prominent' (the author's name is within parentheses):** … (Restouin et al., 2009) … Or **'Author prominent' (the author's name is outside the parentheses):** Restouin et al. (2008) …	Restouin, A., Aresta, S., Prébet, T., Borg, J., Badache, A., & Collette, Y. (2009). A simplified, 96-well–adapted, ATP luminescence–based motility assay. BioTechniques, 47, 871–875. doi: 10.2144/000113250
Journal article with more than 8 or more authors	**Information prominent' (the author's name is within parentheses):** … (Steel et al., 2010).. Or **'Author prominent' (the author's name is outside the parentheses):** Steel et al. (2010) …	Steel, J., Youssef, M., Pfeifer, R., Ramirez, J. M., Probst, C., Sellei, R., ... Pape, H. C. (2010). Health-related quality of life in patients with multiple injuries and traumatic brain injury 10+ years postinjury. *Journal of Trauma: Injury, Infection, and Critical Care,* 69(3), 523-531. doi: 10.1097/ TA.0b013e3181e90c24
Journal article with DOI	**Information prominent' (the author's name is within parentheses):** … (Shaw et al., 2005).. Or **'Author prominent' (the author's name is outside the parentheses):** Shaw et al. (2005) …	Shaw, K., O'Rourke, P., Del Mar, C., & Kenardy, J. (2005). Psychological interventions for overweight or obesity. *The Cochrane database of systematic reviews* (2). doi:10.1002/14651858.CD003818. pub2

Newspapers

Newspapers	**Insertion in Text**	**In Reference List**
Newspaper article – with an author	… (Waterford, 2007)...	Waterford, J. (2007, May 30). Bill of rights gets it wrong. *The Canberra Times*, p. 11.
Newspaper article – without an author	… ("Internet pioneer", 2007) …	Internet pioneer to oversee network redesign. (2007, May 28). *The Canberra Times*, p. 15.
Article in an newsletter	... ("Australians and the Western Front", 2009) ...	Australians and the Western Front. (2009, November). *Ozculture newsletter*. Retrieved from http://www.cultureandrecreation. gov.au/ newsletter/

PERTANIKA JOURNALS

Conference / Seminar Papers

Conference / Seminar Papers	Insertion in Text	In Reference List
Print – *If the paper is from a book, use the Book chapter citation format. If it is from regularly published proceedings (e.g. annual), use the Journal article citation format.*	… (Edge, 1996) … Or Edge (1996) …	Edge, M. (1996). Lifetime prediction: Fact or fancy? In M. S. Koch, T. Padfield, J. S. Johnsen, & U. B. Kejser (Eds.), *Proceedings of the Conference on Research Techniques in Photographic Conservation* (pp. 97-100). Copenhagen, Denmark: Royal Danish Academy of Fine Arts.
Online	… (Tester, 2008) … Or Tester (2008) …	Tester, J. W. (2008). The future of geothermal energy as a major global energy supplier. In H. Gurgenci & A. R. Budd (Eds.), *Proceedings of the Sir Mark Oliphant International Frontiers of Science and Technology Australian Geothermal Energy Conference*, Canberra, Australia: Geoscience Australia. Retrieved from http://www.ga.gov.au/image_cache/GA11825.pdf

Government Publications

Government Publications	Insertion in Text	In Reference List
Government as author	**First in-text reference: Spell out the full name with the abbreviation of the body.** … (Department of Finance and Administration [DOFA], 2006) … **Subsequent in-text reference/s:** **Use the abbreviation of the body.** … (DOFA, 2006)…	Department of Finance and Administration. (2006). Delivering Australian Government services: Managing multiple channels. Canberra, Australia: Author.
Government report - online	**First in-text reference: Spell out the full name with the abbreviation of the body.** … (Department of the Prime Minister and Cabinet [PM&C], 2008) … **Subsequent in-text reference/s:** **Use the abbreviation of the body.** … (PM&C, 2008)…	Department of the Prime Minister and Cabinet. (2008). *Families in Australia: 2008.* Retrieved June 1, 2019, from http://www.....

8. General Guidelines

Abbreviations: Define alphabetically, other than abbreviations that can be used without definition. Words or phrases that are abbreviated in the introduction and following text should be written out in full the first time that they appear in the text, with each abbreviated form in parenthesis. Include the common name or scientific name, or both, of animal and plant materials.

Acknowledgements: Any individuals and entities who have contributed should be acknowledged appropriately.

Authors' Affiliation: The primary affiliation for each author should be the institution where the majority of their work was done. If an author has subsequently moved to another institution, the current address may also be stated in the footer.

Co-Authors: The commonly accepted guideline for authorship is that one must have substantially contributed to the development of the paper and share accountability for the results. Researchers should decide who will be an author and what order they will be listed depending upon their order of importance to the study. Other contributions should be cited in the manuscript's Acknowledgements.

Copyright Permissions: Authors should seek necessary permissions for quotations, artwork, boxes or tables taken from other publications or from other freely available sources on the Internet before submission to Pertanika. Acknowledgement must be given to the original source in the illustration legend, in a table footnote, or at the end of the quotation.

Footnotes: Current addresses of authors if different from heading may be inserted here.

Page Numbering: Every page of the manuscript, including the title page, references and tables should be numbered.

Spelling: The journal uses American or British spelling and authors may follow the latest edition of the Oxford Advanced Learner's Dictionary for British spellings. Each manuscript should follow one type of spelling only.

SUBMISSION OF MANUSCRIPTS

All submissions must be made electronically using the **online submission system ScholarOne™**, a web-based portal by Clarivate Analytics. For more information, go to our web page and click "**Online Submission**".

Submission Checklist

1. **MANUSCRIPT**:

   Ensure your MS has followed the Pertanika style particularly the first four pages as explained earlier. The article should be written in a good academic style and provide an accurate and succinct description of the contents ensuring that grammar and spelling errors have been corrected before submission. It should also not exceed the suggested length.

2. **DECLARATION FORM:**

- Author has to sign a declaration form. In signing the form, authors declare that the work submitted for publication is orginal, previously unpublished, and not under consideration for any publication elsewhere.

- Author has to agree to pay the publishing fee once the paper is accepted for publication in Pertanika.

Note:

COPYRIGHT FORM: Author will be asked to sign a copyright form when the paper is accepted. In signing the form, it is assumed that authors have obtained permission to use any copyrighted or previously published material. All authors must read and agree to the conditions outlined in the form, and must sign the form or agree that the corresponding author can sign on their behalf. Articles cannot be published until a signed form *(original pen-to-paper signature)* has been received.

Visit our Journal's website for more details at http://www.pertanika.upm.edu.my/

ACCESS TO PUBLISHED MATERIALS

Under the Journal's open access initiative, authors can choose to download free material (via PDF link) from any of the journal issues from Pertanika's website. Under "**Browse Journals**" you will see a link, "*Regular Issue*", "*Special Issue*" or "*Archives*". Here you will get access to all current and back-issues from 1978 onwards. No hard copy of journals or off prints are printed.

Visit our Journal's website at
http://www.pertanika.upm.edu.my/regular_issues.php for "Regular Issue"
http://www.pertanika.upm.edu.my/cspecial_issues.php for "Special Issue"
http://www.pertanika.upm.edu.my/journal_archives.php for "Archives"

PUBLICATION CHARGE

Upon acceptance of manuscript, a processing fee of RM 750 / USD 250 will be imposed on authors; RM 750 for corresponding author affiliated to institution in Malaysia; USD 250 for corresponding author affiliated to institution outside Malaysia. Payment must be made online at https://paygate.upm.edu.my/action.do?do=

Any queries may be directed to the **Chief Executive Editor's** office via email to executive_editor.pertanika@upm.edu.my

Hydrological Response of Semi-Arid River Catchment to Rainfall and Temperature Fluctuations 2333
Da'u Abba Umar, Mohammad Firuz Ramli, Ahmad Zaharin Aris and Muhammad Amar Zaudi

Mathematical Sciences

Some Relation Properties of Rough Neutrosophic Multisets 2351
Suriana Alias, Daud Mohamad and Adibah Shuib

Short communication
Characterization of Model on Zoning Process 2373
Ompal Singh

Two Point Block Multistep Methods with Trigonometric–Fitting for Solving Oscillatory Problems 2381
Aini Fadhlina Mansor, Fudziah Ismail and Norazak Senu

Chemical Sciences

Effect of Parboiling on *in vitro* Physiological Antioxidant Capacity of Brown Rice 2399
Elaine Lee, Phene Neoh Pei Nee, Lee Huei Hong, Wong Sie Chuong, Wong Tze Jin, Koo Lee Feng and Yiu Pang Hung

Characteristics of Chitosan Nanoparticles extracted from Sea Cucumber (*Holothuria scabra*) as Source Materials for Glucosamine 2409
Sumarto, Bustari Hasan, Rahman Karnila and Mery Sukmiwati

Material Sciences

Preparation and Characterization of Zeolite type 4A using Kaolin from Ajebo, Nigeria 2427
Henry Ekene Mgbemere, Ikenna Christopher Ekpe, Ganiyu Lawal, Henry Ovri and Anna-Lisa Chaudhary

Physical Sciences

Scattered Radiation Dose Calculation in the Multi-Bay Resuscitation Room using MCNP5 Code 2439
Marianie Musarudin, Ruzainah Mohd Nordin and Nik Kamarullah Ya Ali

Applied Sciences and Technologies

Modelling and Optimization of Torrefied Pellet Fuel Production 2139
Ahmad Hafizi Awang, Abdulhalim Abdulrazik, Azuin Mad Noor and Aainaa Izyan Nafsun

Techniques and Algorithms for Selection of Number and Locations of Temperature Sensors for Greenhouse 2153
Suman Lata and Harish Kumar Verma

Three Phase Grid Connected Neutral Point Clamped (NPC) Multilevel Inverter Fed by Two Wind Turbines 2173
Sanusi Alabi Kamilu and Saad Mekhilef

Efficacy of Silvernano Bacterial Cellulose Insole on *Micrococcus sedentarius* Growth 2189
Kanchana Sitlaothaworn and Manussawee Dechkla

Elemental Concentration in Three Different Fish Species Captured from Oluwa River, Okitipupa, Ondo State, Nigeria 2201
Thompson Faraday Ediagbonya and Ramond Babatunde Adesokan

Fatty Acid Profile and Production of Fatliquor from *Canarium schweinfurthii* Mesocarp Oil 2221
Adachukwu Nwaka Nkwor, Pius Oziri Ukoha, William Robert Wise, Njemuwa Njoku Nwaji and Karl Flowers

Earth Sciences

Assessment of Runoff via Precipitation using Neural Networks: Watershed Modelling for Developing Environment in Arid Region 2245
Sandeep Samantaray, Abinash Sahoo and Dillip Kumar Ghose

Prediction of Groundwater Contaminants from Cattle Farm using Visual MODFLOW 2265
Mohammad Nazri Ebrahim, Hasfalina Che Man, Mohamed Azwan Mohamed Zawawi and Muhammad Hazwan Hamzah

Spatial Pattern and Temporal Variability of Sea Level Anomaly and Geostrophic Current in the Eastern Indian Ocean from Satellite Altimetry 2281
Julian Saputra, Jonson Lumban-Gaol, James Parlindungan Panjaitan and Agus Saleh Atmadipoera

Strength Assessment of Mining Soil Treated with Steel Slag as Liner Material for Retention Pond 2305
Noratiqah Masri, Faradiella Mohd Kusin, Nik Norsyahariati Nik Daud and Hasfalina Che Man

Distribution of Dissolved and Particulate Fe in an Estuarine System at Bagan Pasir, Perak 2317
Syaza Syafiqah Muhamad Kamel, Khairul Nizam Mohamed, Ferdius@Ferdaus Mohamat Yusuff and Abd Muhaimin Amiruddin

Long-Term Water Quality Assessment in a Tropical Monsoon 1955
Jabir Harna Abdulkareem, Da'u Abba Umar, Alhassan Idris Gabasawa, Chinedum Anyika and Nor Rohaizah Jamil

Case study
Monitoring Urban Green Space (UGS) Changes by Using High Resolution Aerial Imagery: A Case Study of Kuala Lumpur, Malaysia 1971
Junainah Abu Kasim, Mohd Johari Mohd Yusof and Helmi Zulhaidi Mohd Shafri

Engineering Sciences

Computational Study of Mixed Heat Convection in Annular Space between Concentric Rotating Inner and Wavy Surface Outer Cylinders 1991
Qusay Rasheed Al-Amir, Farooq Hassan Ali Alinnawi and Qusay Adnan

Strategies to Improve Communication Management within Virtual Project Teams 2015
Nuzul Azam Haron, Law Teik Hua, Salihudin Hassim, Fathollah Eftekhari, Muhammad Tahir Muhammad and Aizul Nahar Harun

Assessment of External Risk Factors on Construction Project Schedule Using Risk Importance Index (RII) 2031
Saiful Husin, Abdullah, Medyan Riza and Mochammad Afifuddin

Detection of Nitrobenzene using a coated Quartz Crystal Microbalance with a Parametric Modeling Approach 2049
Dipali Ramdasi and Rohini Mudhalwadkar

Comparison of Overall Heat Transfer Coefficient between Shell and Tube and Spiral Coil Heat Exchangers 2065
Meseret Nasir Reshid, Girma Tadesse Chala and Wan Mansor Wan Muhamad

Optimization of Glass-beads Water Slurry flow Characteristics using Taguchi's Decision-making Technique coupled with Genetic Algorithm 2079
Shofique Uddin Ahmed, Rajesh Arora and Ranjana Arora

Case study
An Eco-Innovative Green Design Method by QFD and TRIZ Tools - A Case Study of Brass-Ware Manufacturing 2109
Anhad Malhotra, Sonu Rajak and Sanjay Kumar Jha

Utilisation of *Cassia surattensis* Seeds as Natural Adsorbent for Oil Content Removal in Oilfield Produced Water 2123
Mohmad Khalel Ibrahim, Ali Abd Al-Hassan and Ahmed Samir Naje

Time Adaptive Collaborative Filtering for Movie Recommendation 1783
Suganeshwari Gopalswamy and Syed Ibrahim Peer Mohamed

Criminality Recognition using Machine Learning on Malay Language Tweets 1803
Nurul Hashimah Ahamed Hassain Malim, Saravanan Sagadevan and Nurul Izzati Ridzuwan

Measuring Tajweed Augmented Reality-Based Gamification Learning Model (TARGaLM) Implementation for Children in Tajweed Learning 1821
Nurtihah Mohamed Noor, Marina Ismail, Rahmah Lob Yussof and Fakhrul Hazman Yusoff

Environmental Sciences

Review article
Sago Wastes and Its Applications 1841
Nurhusni Amin, Nordin Sabli, Shamsul Izhar and Hiroyuki Yoshida

Optimization of Process parameters for Decolorization of Azo Dye Remazol Golden Yellow by *Bacillus firmus* using Biostatistical Designs 1863
Palanivelan Ramachandran, Subashni Bhoopathy, Ramya Suseenthar and Ayyasamy Pudukkadu Munusamy

Water Quality Effects on the Optimal Water Resources Operation in Great Karun River Basin 1881
Iman Ebrahim Bakhsipoor, Seyed Mohammad Ashrafi and Arash Adib

Evaluation of Aerosol Optical Thickness over Malaysia Based on Multi-Source Ground and Satellite Data 1901
Khaled Ali Ahmed Ben Youssef, Ahmad Makmom Abdullah, Helmi Zuhaidi Mohd Shafri and Zulfa Hanan Ash'aari

Growth Performance of Blood Cockle (Tegillarca granosa) within Kongkong Laut Estuaries, Masai, Johor 1917
Amirul Azuan Md Joni, Ferdaus Mohamat Yusuff, Khairul Nizam Mohamed, Faradiella Mohd Kusin and Syaizwan Zahmir Zulkifli

Spatial and Vertical Metals Distribution in Sediment Cores from Kongkong Laut Estuary, Masai, Johor 1929
Sara Azuana Khalidah Khalid, Ferdaus Mohamat Yusuff, Faradiella Mohd Kusin, Aziz Arshad and Zulfa Hanan Asha'ari

Concentration of Heavy Metals in Street Dust and Surface Soils in Urban and Peri-Urban Regions in the Kuala Lumpur Metropolitan Area 1941
Ahmad Ismail, Munirah Hanapiah, Syaizwan Zahmir Zulkifli and Hishamuddin Omar and Nur Amiera Kamarudin

Monitoring of Fracture Healing Process by Acousto-Ultrasonic Technique 1603
Sharad Shrivastava and Ravi Prakash

MRN Complex and ATM Kinase Inhibitors Impacts towards UVC-Treated Zebrafish Embryonic Development 1615
Shazrul Fazryl, Malina Kumaran, Nahid Khalili, Ibrahim Mahmood, Herryawan Ryadi Eziwar Dyari, Nur Hidayah Jamar, Douglas Law, Nik Marzuki Sidik, Nor Azizan Abdul Murad and Yi Chieh Lim

Identifying Analogues Of 2-Deoxyglucose, Alpha-D-Glucose and Beta-D-Glucose-6-Phosphate as Potential Inhibitors of Human Hexokinase II for the Development of Anti-Dengue Therapeutics 1625
Suriyea Tanbin, Nurhainis Ogu Salim and Fazia Adyani Ahmad Fuad

Information, Computer and Communication Technologies

Review article
Feature Extraction and Classification Techniques of MODI Script Character Recognition 1649
Solley Joseph and Jossy George

Resolution Adaptive Threshold Selection for Gradient Edge Predictor in Lossless Biomedical Image Compression 1671
Urvashi, Meenakshi Sood and Emjee Puthooran

Hybrid Feature Extraction and Machine Learning Approach for Fruits and Vegetable Classification 1693
Nimratveer Kaur Bahia, Rajneesh Rani, Aman Kamboj and Deepti Kakkar

A New Method to Construct 4-Cycle-Free Parity-Check Matrices for Regular LDPC Codes 1709
Djamel Slimani, Adda Ali-Pacha and Naima Hadj-Said

Matching Fingerprint Images for Biometric Authentication using Convolutional Neural Networks 1723
Abdulmawla Najih, Syed Abdul Rahman Al-Haddad Syed Mohamed, Abdul Rahman Ramli Shaiful Jahari Hashim and Nabila Albannai

Improved Salient Object Detection via Boundary Components Affinity 1735
Nur Zulaikhah Nadzri, Mohammad Hamiruce Marhaban, Siti Anom Ahmad, Asnor Juraiza Ishak and Zalhan Mohd Zin

Detection and Recognition via Adaptive Binarization and Fuzzy Clustering 1759
Saad Mohmad Saad Ismail, Siti Norul Huda Sheikh Abdullah and Fariza Fauzi

Pertanika Journal of Science & Technology
Vol. 27 (4) Oct. 2019

Contents

Foreword
Abu Bakar Salleh i

Medical and Health Sciences

Review article
Medical Imaging Literature in MyCite 1471
Nurul Husna Kamarudin, Nor Azlina A. Rahman and Zainul Ibrahim Zainuddin

Review article
Psychosocial Predictors of Adolescent Aggression 1485
Fatimah Ahmad Fauzi, Nor Afiah Mohd Zulkefli, Halimatus Sakdiah Minhat and Norliza Ahmad

Review article
Timing and Prognostic Factors of Tuberculosis Treatment Interruption 1509
Qudsiah Suliman, Salmiah Md. Said and Lim Poh Ying

Simulation Work on Blood Glucose Control for Type 1 Diabetes using Modified Hovorka Equations 1527
Ayub Md Som, Nur Farhana Mohd Yusof, Sherif Abdulbari Ali and Nurafiqah Zawawi

Composite Resin Veneer Systems: An *In Vitro* Study to Evaluate Surface Roughness Changes with Different Surface Treatments and Shear Bond Strength 1539
Nik Zarina Nik Mahmood, Amar Mohammed Thiyab and Mohamed Ibrahim Abu Hassan

Exploring Potential Neuroprotective Properties of Aqueous *Centella asiatica* Extract in Chronic Stress-induced Rats 1561
Muhamad Zulhusni Abdul Wahab, Muhammad Zaim Zainal Abidin, Fatin Nadzirah Zakaria, Mohamad Aris Mohd Moklas, Zulkhairi Amom, Nina Keterina Hashim and Razif Dasiman

Forensic Parameter and Paternity Testing of Tpa-25 Element in Kelantan-Malay and Jawa-Malay 1575
Nurul Syafika Muslimin, Nur Hilwani Ismail and Farida Zuraina Mohd Yusof

Effect of Diet Regime on the Development and Survival of *Aedes albopictus* (Skuse) (Diptera: Culicidae) 1589
Nurul Syafika Muslimin, Nur Hilwani Ismail and Farida Zuraina Mohd Yusof

Pertanika Editorial Office, Journal Division,
Office of Deputy Vice Cahncellor (R&I),
1st Floor, IDEA Tower II,
UPM-MTDC Technology Centre
Universiti Putra Malaysia
43400 UPM Serdang
Selangor Darul Ehsan
Malaysia
http://www.pertanika.upm.edu.my
E-mail: executive_editor.pertanika@upm.edu.my
Tel : +603 9769 1622/1616

PENERBIT UPM
UNIVERSITI PUTRA MALAYSIA
BERILMU BERBAKTI
PRESS

http://penerbit.upm.my
E-mail: penerbit@upm.edu.my
Tel : +603 9769 8851

ISSN 0128-7680
9 770128 768083 10