



2023 PROCEEDING



4th IRCASET 2023

**4TH INTERNATIONAL RESEARCH
CONFERENCE ON APPLIED SCIENCES AND
ENGINEERING TECHNOLOGY 2023
(4TH IRCASET 2023)**

**PROCEEDING:
4TH INTERNATIONAL RESEARCH
CONFERENCE ON APPLIED SCIENCES AND
ENGINEERING TECHNOLOGY 2023
(4TH IRCASET 2023)**



**GLOBAL ACADEMIC EXCELLENCE (M) SDN BHD
2023**

PROCEEDING:



4TH INTERNATIONAL RESEARCH CONFERENCE ON APPLIED SCIENCES AND ENGINEERING TECHNOLOGY 2023 (4TH IRCASET 2023)

27-28 May 2023

Copyright 2023

Global Academic Excellence (M) Sdn Bhd

All rights reserved. No part of this publication may be reproduced in any form or by any means, whether electronically and/or mechanically including photocopying, recording or any means of storage of information and retrieval system, without the written consent of Global Academic Excellence (M) Sdn Bhd / Author.

eISBN: 978- 967-2426-49-3

Published By:



GLOBAL ACADEMIC EXCELLENCE (M) SDN BHD -1257579-U

Lot 2 -11 Arked MARA Kota Bharu, Jalan Dato Pati, 15000 Kota Bharu ,Kelantan, MALAYSIA

Website: www.egax.org

Email: admin@egax.org

TABLE OF CONTENTS

IMAGING OF BUBBLY FLOW USING ULTRASOUND	1
EFFECT OF SLOPE GRADIENT ON THE APPLICATION OF VETIVER GRASS FOR SLOPE STABILIZATION	10
THE PERFORMANCE OF VETIVER ROOT GROWTH WITH EM PB AND COCONUT FIBRE – PHYSICAL MODEL	22
COMPRESSIVE SENSING PATH FOR OPTIMAL DATA TRANSMISSION IN UNDERWATER ACOUSTIC SENSOR NETWORK.....	36
IMPROVED ERROR FREE CLUSTERING IN COGNITIVE RADIO NETWORKS USING MODIFIED PARTICLE SWARM OPTIMISATION	46
BLOCKCHAIN-BASED DEEP LEARNING FOR SUSTAINABLE AGRICULTURAL SUPPLY CHAIN MANAGEMENT	55
SECUREML BASED CLASSIFICATION FOR INTERNET OF THINGS BASED SECURED TRANSACTION OF DATA	65
SOFTWARE MODEL CHECKING DISTRIBUTED APPLICATIONS: A HYBRID APPROACH	76
HEPATITIS C INFECTION IN HEMODIALYSIS PATIENTS, MAKKAH, SAUDI ARABIA.....	93
DEVOPS ADOPTION GUIDELINES, CHALLENGES, AND BENEFITS: A SYSTEMATIC LITERATURE REVIEW.....	109
A REVIEW: CHARACTERIZATION OF ALUMINIUM-METAL MATRIX COMPOSITE REINFORCED WITH ZIRCONIUM DIBORIDE	132
THE RELATIONSHIP BETWEEN BUILDING CONDITION SPECIFICALLY TO INDOOR AIR QUALITY (IAQ) AND SICK BUILDING SYNDROME (SBS) SYMPTOMS AMONG WARD STAFF AT HOSPITAL PAKAR SULTANAH FATIMAH.....	141
REAL TIME PATIENT VITAL MONITORING AND ALARM SYSTEM WITH PREDICTION OF ANOMALIESAND FUTURE CLINICAL EPISODES USING MACHINE LEARNING MODELS	155

INFLUENCE OF TiO ₂ AND Al ₂ O ₃ NANOPARTICLES ADDITION ON THE IMC GROWTH OF SN-3.0AG-0.5CU LEAD-FREE SOLDER AT DIFFERENT REFLOW TEMPERATURE.....	164
COMPARATIVE INVESTIGATION OF RESISTANCE PREDICTION FOR SURFACE COMBATANT SHIP MODEL USING CFD MODELING	172
EXPERIMENTAL AND CFD RESISTANCE VALIDATION OF NAVAL COMBATANT DTMB 5415-51 MODEL	183
DESIGN OF WATER QUALITY MONITORING SYSTEM BASED ON INTERNET OF THINGS TECHNOLOGY	203
A MULTI-CRITERIA ENODE-B PRE-SELECTION AND HANDOVER TRIGGERING SCHEME FOR LTE-A HETEROGENEOUS WIRELESS NETWORK	216
PORE PRESSURE DIFFUSION WAVES TRANSMISSION IN OIL RESERVOIR	228
AN EFFECTIVE AIR POLLUTION PREDICTION MODEL USING MACHINE LEARNING ALGORITHMS	238
SOCIAL MEDIA TUNING ATTENTION MODEL FOR INFORMATION SHARING BEHAVIOUR	247
EFFECT OF MATRIX PERMEABILITY ON A HORIZONTAL SHALE GAS WELL PERFORMANCE USING A FULLY COUPLED FLUID FLOW WITH GEOMECHANICS MODEL	248
POLYETHERSULFONE – POLYITACONIC ACID (PES-PIA) MEMBRANE FOR HIGHLY SELECTIVE DYES REMOVAL	267
FABRICATION AND CHARACTERIZATION OF NANOCOILS CNT/PLA FILAMENT AT DIFFERENT FILLER’S PERCENTAGES AS ELECTROMAGNETIC WAVE ABSORBERS	276
OPTIMIZATION OF SYNTHESIZING CONDITIONS FOR MXENE (Ti ₃ C ₂) PHOTOCATALYST: EFFECT OF LiF: Ti ₃ AlC ₂ MASS RATIO	291
A FRAMEWORK STUDY ON A REAL-TIME OPERATED SYSTEM FOR EMERGENCY MEDICAL SERVICES (EMS) USING NODEMCU PROCESSOR.....	301

RISK ASSESSMENT TO IMPROVE TOLL ROAD BUSINESS RESILIENCE (CASE STUDY: BALI-MANDARA TOLL ROAD)	315
INDUSTRIAL TRIP GENERATION MODEL IN KARAWANG INTERNATIONAL INDUSTRY CITY (KIIC)	326
ANALYSIS OF TRANSPORTATION ACCESSIBILITY IN BOGOR CITY	337
CRYPTOCURRENCY ADOPTION AMONG SAUDI ARABIAN PUBLIC UNIVERSITY STUDENTS: DUAL STRUCTURAL EQUATION MODELLING AND ARTIFICIAL NEURAL NETWORK APPROACH	346
CONTRIBUTION OF PLASTIC WASTE IN RECYCLES CONCRETE AGGREGATE PAVING BLOCK	367
ELEMENTAL ANALYSIS OF PRINTING INK WITH THE APPLICATION OF LASER-INDUCED BREAKDOWN SPECTROSCOPY (LIBS)- A REVIEW	377

IMAGING OF BUBBLY FLOW USING ULTRASOUND

Sallehuddin Ibrahim^{1*}
Mahdi Faramarzi¹
Mohd Amri Md Yunus¹

¹ Department of Control and Mechatronics Engineering, Fakulti Kejuruteraan Elektrik, Universiti Teknologi Malaysia, 81310 Skudai, Johor, Malaysia

Abstract: *The strict industrial and environmental regulations have lead researchers to utilize certain industrial process monitoring, control and instrumentation. Flow measurement is a vital issue in instrumentation and has been applied in the petrochemical, chemical and nuclear power generation industries. The behavior of a single phase flowing fluid, e.g. a liquid through a pipe, is well investigated. However, the characterization of a multi-component flowing mixture, such as oil and gas, is more difficult and conventional flow meters do not usually give satisfactory measurement results. In this regard, process tomography utilizing non-intrusive sensors is gaining considerable significance in industrial measurements because it does not disrupt the flow process. An ultrasonic tomography system is one of these techniques which is applicable for bubbly flow with high acoustic impedance difference.*

.

Keywords: *Bubbles, Flow, Imaging, Ultrasound*

1. Introduction

Multiphase flows occur in the form of two or more thermodynamic phases in a flow pipe. Multiphase flows can be observed in various natural phenomena and in various industries. Such phases involve gas, liquid or solid and liquids or solids which cannot mix [1]. Two phase flows are the most common multiphase flow where two phases which cannot mix but can coexist in a thermodynamic equilibrium. Bubbly flow is a form of a two-phase flow which occurs at large liquid flows having a small amount of gas bubbles which are dispersed. Bubbly flow can be found in various industries such boiling water nuclear reactors, steam generators and air condition. Research on the operation and transport phenomena of bubble columns have led to better comprehension of the hydrodynamic properties, heat and mass transfer mechanisms as well as the flow regime characteristics [2, 3]. Because of the strict regulations regarding precise flow control mainly in the case of two-phase fluid flow, there is an urgency to develop simpler and precise instrumentations. As such tomography systems can play a significant role in industrial process [4].

Tomography is a method for obtaining images of the internal characteristic of multiphase flow. It has been applied in multifarious industries as it can effectively obtain cross-sectional profiles of a flow and it also has the ability of distinguishing between the components of a dissimilar phase from the continuous one [3]. From these images important data of a process can be extracted which are useful in monitoring a process. From this data various parameters can be obtained such as the void fraction, flow regime and mean velocity [5]. By processing signals

at various positions vital data is obtained which enable investigators to get a complete picture of the region of interest.

During the last five decades there have been extensive investigation to develop tomography systems which can be applied in the industries. As such investigators came out with various tomography systems such as electrical capacitance tomography (ECT), electrical resistivity tomography (ERT), electrical impedance tomography (EIT), optical tomography (OPT), gamma-ray, X-ray, terahertz tomography, positron emission tomography and ultrasonic tomography (UT) [6]. These methods can be categorized as hard-field and soft-field. In the case of the hard field method regardless of medium, the direction of travel of the energy waves from the transmitter is constant. Ultrasonic tomography and magnetic resonance imaging are examples of hard-field tomography. In the case of soft-field, the electric current is introduced to the medium being profiled and an electric field distribution is gained based on the physical electrical properties of that material, enabling a distribution of parameters such as resistance, capacitance or impedance to be reconstructed by a computer to form the tomographic images. Soft field is complex than hard field and as such the formed requires more computer analysis and algorithms to form the image [7]. This is due to the fact that soft field is a nonlinear and on the contrary hard field is linear.

Ultrasound can detect fluctuations in acoustic impedance which is a measure of how much resistance the ultrasound beam experience as it propagates through an object [8]. This characteristics of ultrasound enables it to be applied in liquid/gas two-phase flow having high-acoustic impedance mixtures e.g. bubbly flow [9]. Besides, ultrasound has the advantage of being inexpensive and safe in comparison with radiation-based methods [10].

Ultrasonic tomography combines hardware and software. The hardware is in the form of sensors which are placed at the exterior of a vessel. The circuits include signal conditioning, controllers and a personal computer for processing and monitoring the data. The ultrasonic sensor is excited with an electrical pulse and subsequently ultrasonic signals will be transmitted from the sensor towards the region of interest. There will be scattering and attenuation of the waves inside the pipe when they propagated through the interface of two distinctive materials. Data from all receivers and then they are transmitted to a computer which enable the images to be reconstructed. These images illustrate the distribution of the materials within the pipe.

Image reconstruction algorithms are processed by software which plays an important role in the final stage of cross-sectional monitoring of a pipe. Image reconstruction comprises two parts which is the forward problem and inverse problem. The forward problem is concerned with the output of each sensor and the sensing area utilizing the sensitivity maps whereas the aim of the inverse problem is to reconstruct an image to determine the profile of materials such as gas bubbles inside water.

Two main categories in the image reconstruction methods are the analytical/single step and series expansion/iterative methods [11]. Besides there are some heuristic methods which have been used for image reconstruction including non-linear, artificial neural network [12,13] and fusion methods [14, 15], in which a dual mode tomography is implemented. Contrary to the speed and simple implementation of analytical methods, they have disadvantages in that less number of sensors are utilized or few view data, which leads to less accuracy However, iterative

methods are insensitive to noise and they can reconstruct an optimal image even in the situation where the data is incomplete but they suffered mostly from low computational speed.

2. Ultrasonic Propagation

It is not easy to mathematically model the propagation of ultrasound and its characteristics but in bubbly flow it can be simplified to some simplified formulae. Acoustic impedance is one of the most important parameters that must be taken into consideration when considering the interaction between an ultrasound wave with a material (Z). The acoustic impedance can be mathematically expressed as:

$$Z = \rho c \quad (1)$$

in which ρ is the density of the material and c is the speed of sound inside a material.

When the ultrasonic wave interact with the boundary of two different materials, the wave will be reflected, transmitted or refracted. Reflection and transmission coefficients, R and T are expressed as:

$$R = \left(\frac{Z_2 - Z_1}{Z_2 + Z_1} \right) \quad (2)$$

$$T = \left(\frac{2Z_2}{Z_2 + Z_1} \right) \quad (3)$$

where Z_1 and Z_2 are the external and internal acoustic impedance respectively.

Attenuation caused the amplitude and intensity of ultrasound to be reduced as they travel. It is vital to take into account of attenuation as a decreased amplitude affect the quality of the image.

Attenuation can be described by the following equation

$$I_R(z) = I_T \exp(-\mu z) \quad (4)$$

where I_R is the ultrasonic amplitude at distance z from source, I_T refers to the ultrasonic initial amplitude at the source, z represents the total pass length and μ is the attenuation coefficient in the medium.

The three modes of ultrasonic wave propagation are transmission, reflection and diffraction mode. In theory all three modes can be treated as individual solutions (based on different assumptions) of the same inverse problem of a scattered field produced by an object, or group of objects. If the data obtained by the transducers is about amplitude or transmission time (time of flight) of forward scattering signal, and the reconstruction method is based on straight-line propagation then, the system works as a transmission mode. In the case of reflection-mode technique, the measurement is the same as transmission mode but it is applied to back-scattering signals. As for the transmission and the reflection mode if the reconstruction algorithm is based on the inverse solution of ultrasonic wave equation, it is referred as diffraction mode.

The major scope of this investigation is on the image reconstruction method based on the transmission mode and there are two types of reconstructive ultrasonic tomography based on the transmission mode. The first type is related to attenuation measurements which resulted in the spatial distribution of the acoustic absorption coefficient in the object being measured. The second type utilized the time-of-flight which reproduces the acoustic refractive index. In this research, the first type is used to achieve the research objectives.

Due to the dynamic behavior of flow in industrial pipes a real time system is required to monitoring such process. Ultrasonic wave propagation inside a tomography system is slower compared to other types of tomography systems such as electrical or optical tomography

systems. In this transmission mode ultrasonic system which composed of 16 transducers located around a pipe full of water having a diameter of 110mm, the minimum data acquisition time required for a frame of image is 8.4 ms. Besides the data acquisition time, processing of data and reconstruction of image will increase the total time of completing one frame. As such results from experiment would indicate the delay time length. The measured acquisition time in experiment determines the number of image frames in a second.

3. Image Reconstruction

Reconstruction of images involved detectors collecting cross section projections at different angles around the object which are then combined to produce either a two dimensional or three dimensional images. A suitable algorithm is utilized to reconstruct these images using these projections. The image reconstruction can be reformulated as an inverse problem. For the inverse problem, the exact solution to the reconstruction procedure is difficult to be obtained because of insufficient data and unwanted noise disturbances. Radon formulated a mathematical technique which enabled an object mapped to lower dimensions employing projections that could be reconstructed to the original. The solution to this mathematical problem has been implemented for various applications such as astronomy and optics.

In the mathematical aspect, Radon transform can be expressed as the integral transform of a function over straight lines. The following equation describes a one dimensional projection of a cross section of an object $f(x, y)$:

$$p_{\theta}(t) = \int_{-\infty}^{\infty} f(x, y) ds \quad (5)$$

where $p_{\theta}(t)$ represents the projections of $f(x, y)$, θ is the angle of rotation, t and s are coordinates in Radon space and the integral of $f(x, y)$ is along the line s .

The coordinate rotation matrix is:

$$\begin{bmatrix} t \\ s \end{bmatrix} = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad (6)$$

The mathematical expression for the inverse radon transform is:

$$f(x, y) = R^{-1}\{p_{\theta}(t)\} = \frac{1}{2\pi^2} \int_0^{\pi} \int_{-\infty}^{\infty} \frac{1}{r \cos(\theta - \varphi) - t} \frac{\partial p_{\theta}(t)}{\partial t} d\theta dt \quad (7)$$

where r and φ are polar coordinates and θ is the angle of rotation from 0 to π .

This represents radon inversion which is needed for all measurement data from 0 to π . Practically, to implement radon transform the number of sensors is not infinite. The level of accuracy varies with different sensor geometries and various ways of discretization.

Image reconstruction involved the forward problem and the inverse problem. The forward problem involved simulation of a sensitivity map having a discrete line from the output of the transmitter to the input of the receiver. A sensitivity map consists of a matrix having all zeros with the exception of the elements forming a part of the line from the transmitter to the receiver. It is good to have an accurate forward model so as to enhance image reconstruction but there is no accurate model and as such an estimated method is utilized.

Linear back projection or simply known as LBP is the simplest reconstruction procedure which provides an estimation of the density at a point by summing (integrating) all the projections that pass through it at various angles therefore. It is assumed in LBP is that the projections are ideal straight lines and therefore it produces an image with smooth details.

In fan-beam projection geometry, the LBP method computes the density of objects in each pixel of reconstructed image based on the following equation:

$$f(x, y) = \int_0^{2\pi} p_\theta(t) d\theta \quad (8)$$

in which $f(x, y)$ is the object, $p_\theta(t)$ is the radon space function and θ is the angles between projection and x axis.

In parallel beam geometry the equation can be modified as:

$$f(x, y) = \int_0^\pi p_\theta(t) d\theta \quad (9)$$

This is due to the different geometry between parallel and fan beam projections. In fan beam the scan range for sensors is from 0 to 2π to cover all ROI while in parallel beam half of angles range is enough.

The flowchart in Figure 1 illustrates the steps used for the image reconstruction algorithm.

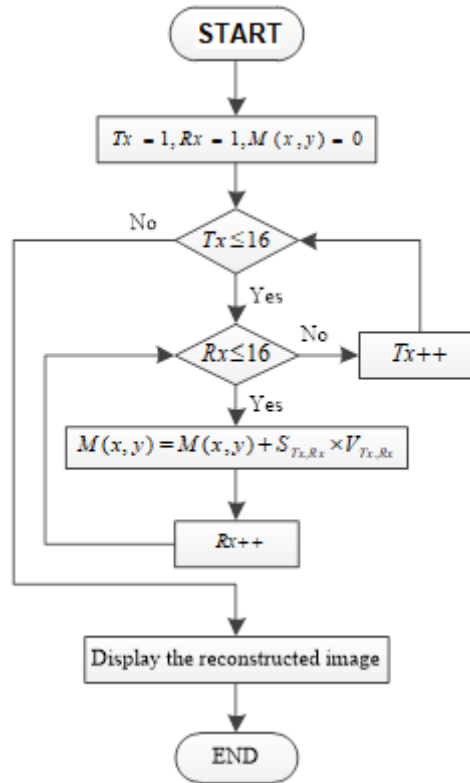


Figure 1: LBP Algorithm Flowchart

As shown in Figure 1 the program starts with initializing Tx (transmitter number), Rx (receiver number) and M (image matrix) to 1,1 and 0, respectively. It means that all pixels in the M matrix set to zero. Then in each loop one projection multiplied by the sensitivity map ($S_{Tx,Rx}$) and the resulted matrix is added to the image matrix (M) using the LBP equation. Finally the image matrix is displayed by the monitor.

Several phantoms representing various bubbly flow regimes were used to evaluate the performance of different reconstruction methods. Phantom 1 is a single bubble with a cross-sectional area of 48mm^2 whereas phantom 2 is a single bubble having a cross-sectional area of

192mm². Phantom 3 and 5 are multiple bubbles with the same size but at different locations. Phantom 4 consists of a small bubble and a larger bubble of different sizes. The results were obtained in static mode and low voidage bubbly flow was assumed.

The ultrasonic system is shown in Figure 2. The pipe is an acrylic transparent pipe which enables visual observations to be made and sixteen ultrasonic transceivers which were mounted on a circular array around the pipe. The pipe has an inner diameter of 100 mm and an outer diameter of 110mm. When the ultrasonic waves is transmitted it experience a loss of transmission power in the air. Hence, the surface of ultrasonic sensors is covered by glycerin to prevent air bubbles being trapped between the sensors surface and the pipe wall. In order to create phantom, the phantoms are placed inside the PVC pipe with different diameters of 5mm, 10mm and 20mm. As part of the noise reduction strategy, coaxial cables were utilized to connect sensors to the main board. The main board includes the transmitter unit, the receiver units and the controller unit. The transmitter unit consists of a MOSFET which generates the ultrasonic burst tones. Each receiver unit consists of two amplifiers, a diode for rectifying received signals, a peak detector and sample and hold IC to keep the peak of the received signal. The controller unit was used to send control signals and collect the output of all receivers and send these data to the computer.

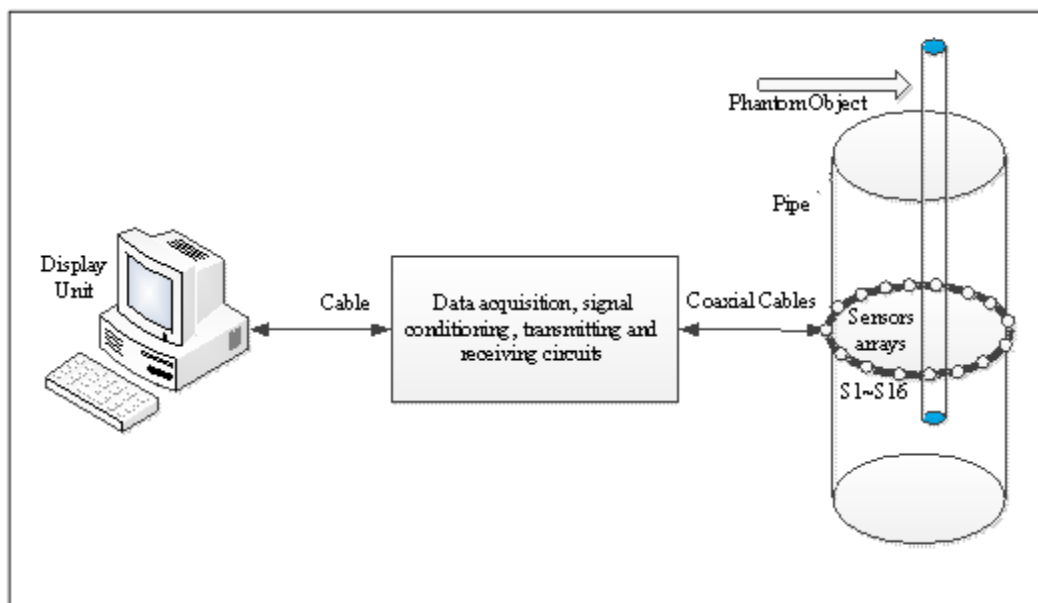


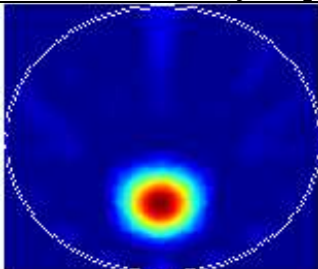
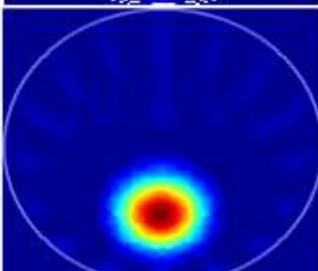
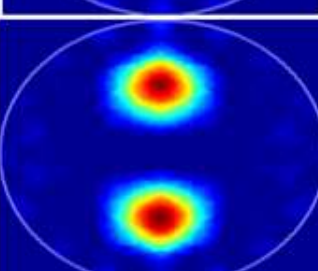
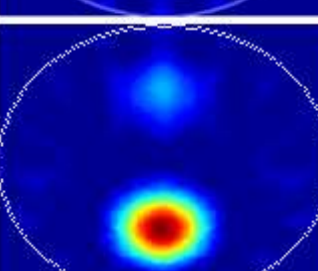
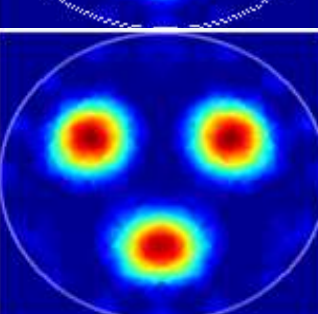
Figure 2: Ultrasonic Tomography Measurement Setup

4. Results and Discussion

Table 1 illustrates the results of various images reconstructed using the LBP algorithm showing different locations of phantoms placed inside a vertical pipe. Table 1 shows that the images reconstructed by LBP method are smeared by noise which occurred due to the limitation of the LBP method. Noise can occur at the sensor response level when a receiver is activated to collect the data or during signal conditioning process which composed of electrical circuits. From the reconstructed images of phantoms in Table 1, the location of the bubbles can be discerned inside the pipe. Phantom 1 shows a single bubble whereas Phantom 2 shows a single bubble of a larger size. Phantom 3 shows two bubbles at different locations. Phantom 4 shows the different conditions in which the size of bubbles is not same. Phantom 5 shows three bubbles

at different locations. The system has successfully show the locations of the bubbles inside the pipe.

Table 1: Images of Phantoms Using the LBP Method

Phantom	Normal sensitivity map
Phantom 1	
Phantom 2	
Phantom 3	
Phantom 4	
Phantom 5	

5. Conclusions

Stringent rules for industries and the environment have led to various research on industrial process monitoring, control and instrumentation. Flow measurement is extremely important in the process instrumentation and has been implemented in various industries such as the petrochemical, chemical and nuclear power generation industries. However, the characterization of a multi-component flowing mixture, such as oil and gas, is more difficult and conventional flow meters do not usually yield satisfactory measurement results. As such, process tomography using non-intrusive sensors is gaining considerable significance in industrial measurements since measurement within the pipe might disrupt the flow process. The ultrasonic system is one of these techniques which is applicable for two phase flow with high acoustic impedance difference. The characteristics of bubbles in a pipe was investigated in this research. The LBP algorithm can identify the location of bubbles in the pipe whether it is a single bubble or multiple bubbles.

Acknowledgement

The authors would like to acknowledge the assistance and the financial grants from Universiti Teknologi Malaysia and the Ministry of Education (UTMFR Q.J130000.2551.21H06) that help us to perform this research.

References

- [1] Da Silva, M. J. (2008) "Impedance sensors for fast multiphase flow measurement and imaging".
- [2] Sumit R. Hazare, Chinmay S. Patil, Shivam V. Vala, Aniruddha J. Joshi, Jyeshtharaj B. Joshi, Vivek S. Vitankar, Ashwin W. Patwardhan "Predictive analysis of gas hold-up in bubble column using machine learning methods" *Chemical Engineering Research and Design* 184 (2022): 724-739.
- [3] Neha Agarwal, Moonyong Lee, Hyunsung Kim "A Non-Invasive Method for Measuring Bubble Column Hydrodynamics Based on an Image Analysis Technique" *Processes* (2022) 10 no. 8, :1660.
- [4] Wahab, Y. A., R. A. Rahim, M. H. F. Rahiman, S. R. Aw, F. R. M. Yunus, C. L. Goh, *et al.*, "Non-invasive Process Tomography in Chemical MIXTURES-A Review," *Sensors and Actuators B: Chemical* 210 (2015): 602 – 617.
- [5] Hua L., W. Mi, W. Ying-xiang, M. Yi-xin, and W. Richard, "Measurement of oil volume fraction and velocity distributions in vertical oil-in-water flows using ERT and a local probe," *Journal of Zhejiang University Science A*, 6, (2005): 1412-1415.
- [6] Wahab Y. A., R. A. Rahim, M. H. F. Rahiman, H. A. Rahim, S. R. Aw and Juliza Jamaludin, "A Review of Process Tomography Application in Inspection System," *Jurnal Teknologi*, 70 (2014).
- [7] Wei H. and M. Soleimani, "Electromagnetic Tomography for Medical and Industrial Applications: Challenges and Opportunities" *Proceedings of the IEEE*, 101, (2013): 559 – 565.
- [8] Rahiman M.F., Abdul Rahim R., Wei Nyap N., and Kok San C., "Monitoring liquid/gas flow using ultrasonic tomography," *Jurnal Teknologi* 40 (2004).
- [9] Rahiman M. F., R. A. Rahim, and Z. Zakaria, "Design and modelling of ultrasonic tomography for two-component high-acoustic impedance mixture," *Sensors and Actuators A: Physical*, 147, (2008): 409 – 414.
- [10] Steiner G. and Podd F., "A non-invasive and non-intrusive ultrasonic transducer array for process tomography," *Interaction*, 99, (2006): 10-15.

- [11] G. Zeng, "Image reconstruction—a tutorial," *Computerized Medical Imaging and Graphics*, 25, (2001): 97 – 103.
- [12] Cierniak, R., "New neural network algorithm for image reconstruction from fan-beam projections," *Neurocomputing*, 72, (2009): 3238 – 3244.
- [13] Cierniak, R., "A 2D approach to tomographic image reconstruction using a Hopfield-type neural network," *Artificial intelligence in medicine*, 43, (2008): 113-125.
- [14] Shouxiao Li, H. Wang, L. Zhang, and W. Fan, "Image reconstruction of electrical resistance tomography based on image fusion," in *Instrumentation and Measurement Technology Conference (I2MTC), IEEE*, (2011): 1-5.
- [15] Hengrong Lan, Changchun Yang and Fei Gao "A jointed feature fusion framework for photoacoustic image reconstruction" *Photoacoustics* 29, (2023): pp 1-12.

EFFECT OF SLOPE GRADIENT ON THE APPLICATION OF VETIVER GRASS FOR SLOPE STABILIZATION

Azim Mustafa^{1*}
Hasif Zulkafli²
Noorasyikin Mohammad Noh³

^{1,2,3} Faculty of Civil Engineering and Built Environment, Universiti Tun Hussein Onn Malaysia, 86400 Parit Raja, Batu Pahat, Johor, Malaysia

Abstract: *This research describes the unique characteristics of Vetiver Grass, Vetiveria Zizanioides Nash, as a bioengineering technique to use for stabilizing slopes. The application of vegetation as part of slope stability and to control erosion has proof to be cost-effective for stabilising surface of the slope. The objectives of this research are to determine the physical properties of soil, identify root properties of Vetiver Grass and measure the mechanical properties and tensile shear strength of root with soil with different slope gradient (45°, 50° and 60°). The methodologies of this study carried out through modelling and lab work to analyses the objective. Root diameter, root length and root morphology observed to identify root properties of Vetiver Grass. Shear box applied to measure the mechanical properties and tensile strength of root with soil with different slope gradient. This study can be made as a reference to produce slope stabilization method by using Vetiver Grass as it medium of bioengineering technique in the future. As we know, slope stability issues become one of the main problems in construction industry that has been acknowledged as one of the most frequent natural disasters that can lead to great loss in property and life. So, it would be one of the best solutions in order to avoid the slope failure from happened again in the future.*

Keywords: *Vetiver Grass, Slope Stabilization, Root Enhancement, Physical Properties*

1. Introduction

According to [10] Malaysia has experienced several landslides throughout the past. As a consequence of these factors, slope failures often occur gradually. Every year during the monsoon seasons, the occurrence of landslides is common in Malaysia. These landslides either cause closure of roads, affect buildings or worse it sometimes causes casualties. A significant portion of rainfall intensity infiltrates the slope surface, depending on the soil permeability [10].

According to [10], the Public Works Department (JKR) routinely provides vegetative cover on these cut slopes to decrease erosion and to preserve susceptible slopes. Consequently, understanding the soil- root interaction is critical for slope monitoring with vegetative coverings. Both soil and vegetation features and attributes should be thoroughly investigated in order to maximize protection and maintenance efforts. Vegetation in the Malaysian environment has the ability to give immediate mechanical shear strength for slope remediation and long-term advantages [10]. Because of its benefits, the bioengineering approach of plant cover was often used to mitigate slope collapse. The approach is less expensive, grows quickly, and is simple to plant [9].

Even after the slope has been covered with these grasses, slope failures continue to occur. The aim of this study is to better understand the interaction between soil and root when it comes to improving the strength of slopes with different slope gradients. The parametric analysis is used to measure the efficiency of the vegetative cover in protecting the slope. Additionally, the slope geometry (45°, 50° and 60° slope angle) is analysed in better detail to provide for the impact of construction. The applicable plant is Vetiver Grass, and its characteristics and characteristics are underlined. This study emphasises the need of selecting an appropriate vegetation type and researching its properties for slope preservation and protection.

2. Slope Gradient on the Application of Vetiver Grass for Slope Stabilization

A slope was a section of land that was part of a mountain or hill and had to be angled so that it raised more at one end than the other. According to [5], a structure of soil with definite angles was known as slope. The matters that make the slope more at risk of landslides or slope failure was the steep slope as well as abundant and continuous rainfall can cause an increase in surface runoff and groundwater flow. Vegetation as a slope cover has several advantages, including environmental friendliness, ease of maintenance, self-sustainability, and a high degree of biodiversity [12]. According to [3], the benefits of plants include self-repairing, regenerative, and adaptability. This will make it easier to manage the vegetation on the slopes. Because vegetation adapts so well, the root structure will extend to fit in with the ground dirt, which serves as the habitat for the plant. The Vetiver System is a developing soil conservation technique and, more recently, a bioengineering tool. Effective application of the Vetiver System, which has major engineering design and construction requirements, necessitates an understanding of every sector [11].

The tap and fibrous roots systems are the two basic types of roots. In plant families, the tap root is the main root that grows from the radicle. The tap root system persists as the primary root (tap root) from which lateral roots emerge. Fibrous root system – This root system has a cluster of short, fiberlike roots that sprout from the base of the radicle and plumule. Because these roots are sparingly branched, shallow, and horizontally extending, it unable to provide significant anchoring. The shear strength of soil with Vetiver roots is 78 percent higher than soil without Vetiver roots [6]. The capacity of Vetiver Grass to improve the factor of safety of embankment slopes against natural stresses has been demonstrated. According to [7], the morphological qualities of the root system and the tensile strength of individual roots have the greatest impact on stabilising sloping terrain via soil reinforcement.

According to [8], slope gradient has a significant impact on the amount of surface water runoff and soil sediment loss. When the slope angle exceeds a critical threshold, the rate of soil erosion increases logarithmically. In different climatic zones, the slope gradient can have a varying effect, primarily dependent on annual rainfall. According to [2], several of the studies focus on how soil properties changed as a result of the presence of vegetative roots. Direct shear testing, for example, was utilised to investigate and determine additional cohesion caused by vegetation reinforcement, and triaxial tests were performed to evaluate both additional cohesion and increased friction angle caused by root reinforcement [4].

3. Methodology

All the experiments conducted in Geotechnical Engineering Laboratory, FKAAB, UTHM according to the relevant standard. The experiments conducted for physical determination are as follows: Shear box and Root tensile Strength Test. The Shear Box Test and Root Tensile

Strength Test conducted to identify the mechanical properties and tensile shear strength of the root-soil. For the root morphology of Vetiver grass, it was observed by naked eyes. The soil modelling was designed in a rectangular shape with varied slope gradients. This modelling soil dimensions 50 cm x 50 cm x 80 cm shown in Figure 1. This soil modelling box is built by plywood and solid wood. Three models utilised in this work to illustrate the varied slope gradients of 45°, 50°, and 60°. The slope gradient angle used to place the soil into the soil structure modelling. The soil is put into soil modelling based on the 45°, 50° and 60° slope gradients that have been specified. The Vetiver Grass planted for 2 phases of time which is 21 days and 30 days inside the modelling box. The scope of works were then proceed with determination of root morphology and mechanical properties of the root using soil with varied slope gradients.



Figure 1: Physical Root- Soil Modelling with Varies of Slope Gradients

4. Results and Discussion

4.1 Root Morphology

The root morphology was determined by using naked eyes. Figure 2 shows the root morphology for Vetiver grass after extracting out from the soil slope sample. It can be seen that the root morphology can be categorized as fibrous root system.



Figure 2: Root morphology of Vetiver grass

4.2 Direct Shear Strength Test

An experimental test was done with varied vertical loaded was applied it is 10, 20, and 30 kN/m². That loads were selected for investigating the association between shear strength and the features of the soil sample. The value of optimal moisture content, maximum dry density, and dry unit weight of the soil sample was employed in the direct shear test based on the outcome of the standard proctor compaction test. Table 5 shows the results of normal stress and shear stress of soil for slope gradient 45°.

Table 5: Result of Normal Stress and Shear Stress of soil for slope gradient 45°

Normal Stress (kN/m ²)	Shear Stress (kN/m ²)
2.78	7.66
5.56	10.57
8.533	14.95

In order to obtain the angle of friction, a graph of shear stress against normal stress was plotted. The shear stress was calculated from the results. The data has been recorded and tabulated in Table 4.5. The graph has been constructed as shown in Figure 3 and the angle of friction for soil with slope gradient of 45° was determined.

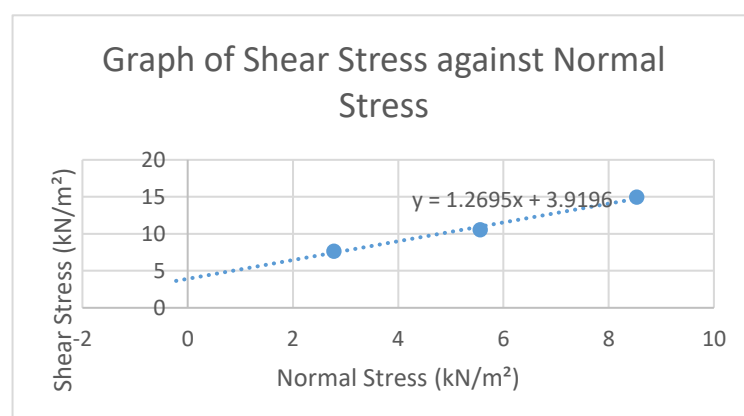


Figure 3: Graph of Shear Stress against Normal Stress

From the graph, the interception of the straight line does not pass through zero. Therefore, the cohesion. $c = 3.92$ kN/m². Besides, the angle of friction is obtained from the plot. From the calculation, the angle of friction of the soil sample is 51.77°. Table 6 shows the results of normal stress and shear stress of soil for slope gradient 50°.

Table 6: Result of Normal Stress and Shear Stress of soil for slope gradient 50°

Normal Stress (kN/m ²)	Shear Stress (kN/m ²)
2.78	6.14
5.56	8.57
8.533	10.21

In order to obtain the angle of friction, a graph of shear stress against normal stress was plotted. The shear stress was calculated from the results. The data has been recorded and tabulated in Table 6. The graph has been constructed as shown in Figure 4 and the angle of friction for soil with slope gradient of 50° was determined.

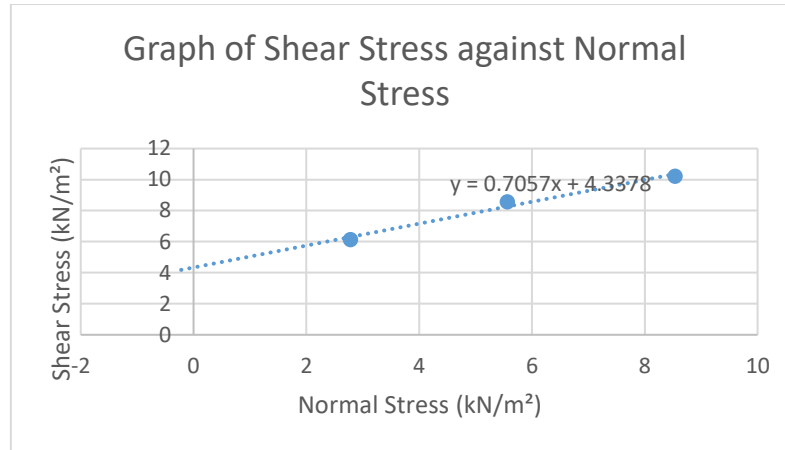


Figure 4: Graph of Shear Stress against Normal Stress

From the graph, the interception of the straight line does not pass through zero. Therefore, the cohesion, $c = 4.34 \text{ kN/m}^2$. Besides, the angle of friction is obtained from the plot. From the calculation, the angle of friction of the soil sample is 35.22°. Table 7 shows the results of normal stress and shear stress of soil for slope gradient 60°.

Table 7: Result of Normal Stress and Shear Stress of soil for slope gradient 60°

Normal Stress (kN/m ²)	Shear Stress (kN/m ²)
2.78	7.29
5.56	8.87
8.533	9.48

In order to obtain the angle of friction, a graph of shear stress against normal stress was plotted. The shear stress was calculated from the results. The data has been recorded and tabulated in Table 7. The graph has been constructed as shown in Figure 5 and the angle of friction for soil with slope gradient of 60° was determined.

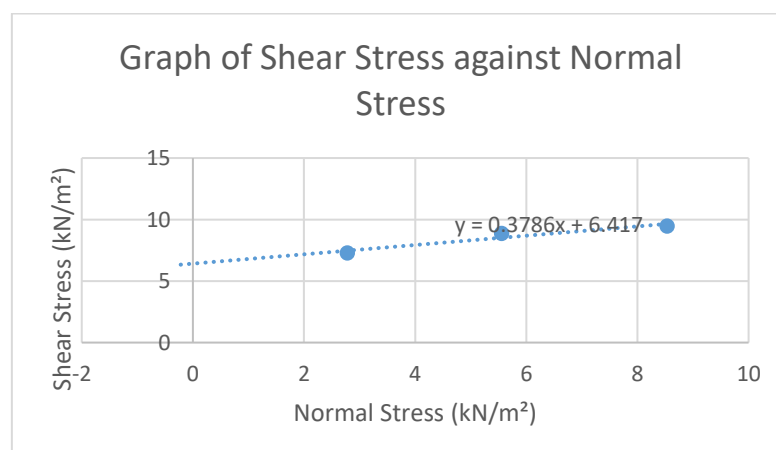


Figure 5: Graph of Shear Stress against Normal Stress

From the graph, the interception of the straight line does not pass through zero. Therefore, the cohesion. $c = 6.42 \text{ kN/m}^2$. Besides, the angle of friction is obtained from the plot. From the calculation, the angle of friction of the soil sample is 20.74° . Table 8 shows the results of normal stress and shear stress of soil with root after 21 days for slope gradient 45° .

Table 8: Result of Normal Stress and Shear Stress of soil with root after 21 days for slope gradient 45°

Normal Stress (kN/m^2)	Shear Stress (kN/m^2)
2.78	11.67
5.56	12.03
8.533	13.73

In order to obtain the angle of friction, a graph of shear stress against normal stress was plotted. The shear stress was calculated from the results. The data has been recorded and tabulated in Table 8. The graph has been constructed as shown in Figure 6 and the angle of friction for soil with root after 21 days for slope gradient 45° was determined.

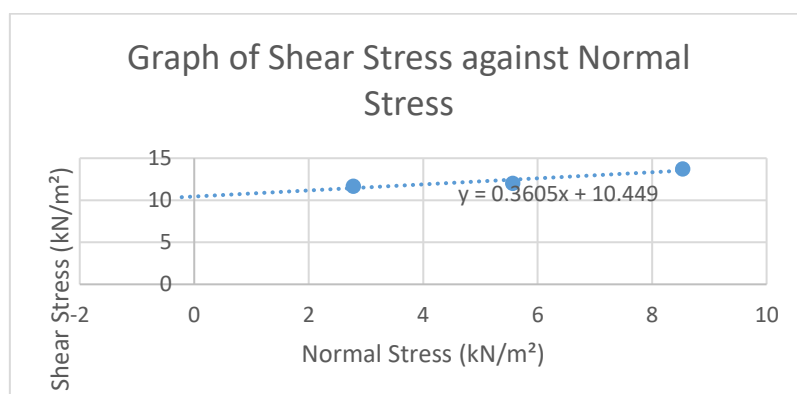


Figure 6: Graph of Shear Stress against Normal Stress

From the graph, the interception of the straight line does not pass through zero. Therefore, the cohesion. $c = 10.45 \text{ kN/m}^2$. Besides, the angle of friction is obtained from the plot. From the calculation, the angle of friction of the soil sample is 19.82° . Table 9 shows the results of normal stress and shear stress of soil with root after 21 days for slope gradient 50° .

Table 9: Result of Normal Stress and Shear Stress of soil with root after 21 days for slope gradient 50°

Normal Stress (kN/m^2)	Shear Stress (kN/m^2)
2.78	11.55
5.56	11.85
8.533	13.73

In order to obtain the angle of friction, a graph of shear stress against normal stress was plotted. The shear stress was calculated from the results. The data has been recorded and tabulated in Table 9. The graph has been constructed as shown in Figure 7 and the angle of friction for soil with root after 21 days for slope gradient 50° was determined.

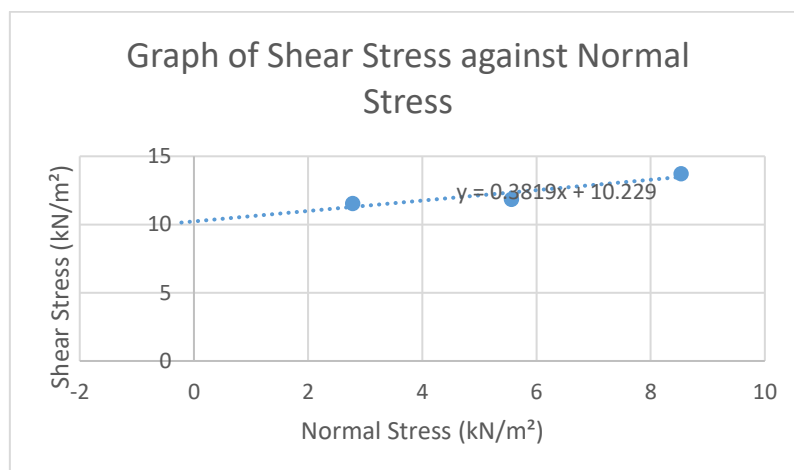


Figure 7: Graph of Shear Stress against Normal Stress

From the graph, the interception of the straight line does not pass through zero. Therefore, the cohesion. $c = 10.23 \text{ kN/m}^2$. Besides, the angle of friction is obtained from the plot. From the calculation, the angle of friction of the soil sample is 20.90° . Table 10 shows the results of normal stress and shear stress of soil with root after 21 days for slope gradient 60° .

Table 10: Result of Normal Stress and Shear Stress of soil with root after 21 days for slope gradient 60°

Normal Stress (kN/m^2)	Shear Stress (kN/m^2)
2.78	12.52
5.56	12.88
8.533	13.25

In order to obtain the angle of friction, a graph of shear stress against normal stress was plotted. The shear stress was calculated from the results. The data has been recorded and tabulated in Table 10. The graph has been constructed as shown in Figure 8 and the angle of friction for soil with root after 21 days for slope gradient 60° was determined.

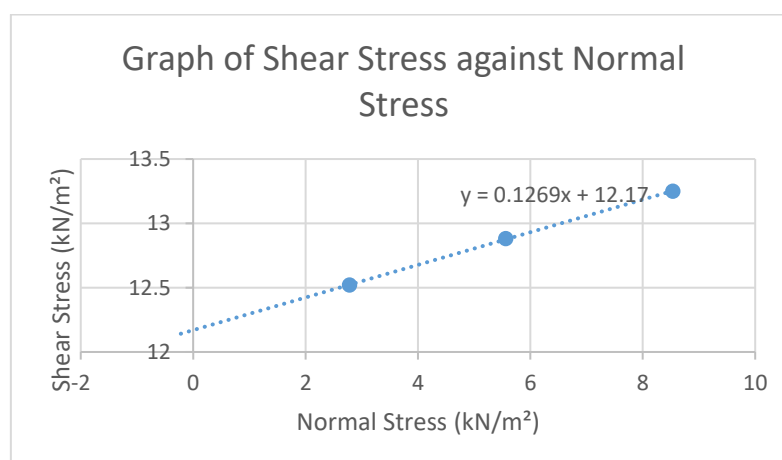


Figure 8: Graph of Shear Stress against Normal Stress

From the graph, the interception of the straight line does not pass through zero. Therefore, the cohesion. $c = 12.17 \text{ kN/m}^2$. Besides, the angle of friction is obtained from the plot. From the calculation, the angle of friction of the soil sample is 7.23° . Table 11 shows the results of normal stress and shear stress of soil with root after 30 days for slope gradient 45° .

Table 11: Result of Normal Stress and Shear Stress of soil with root after 30 days for slope gradient 45°

Normal Stress (kN/m^2)	Shear Stress (kN/m^2)
2.78	10.82
5.56	28.56
8.533	32.93

In order to obtain the angle of friction, a graph of shear stress against normal stress was plotted. The shear stress was calculated from the results. The data has been recorded and tabulated in Table 11. The graph has been constructed as shown in Figure 9 and the angle of friction for soil with root after 30 days for slope gradient 45° was determined.

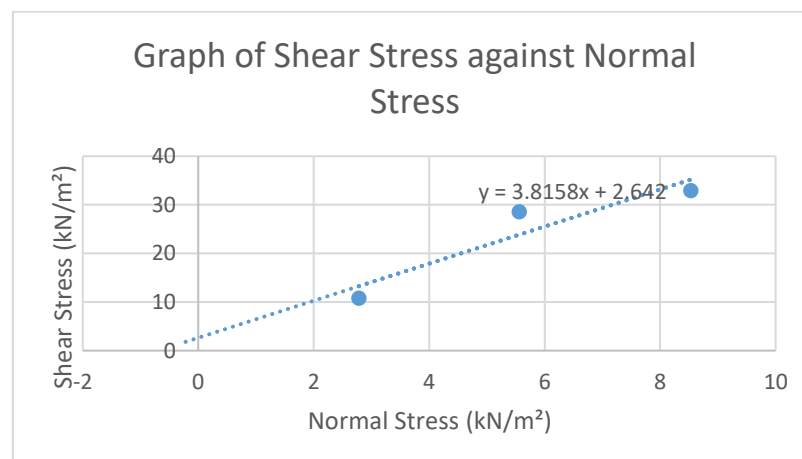


Figure 9: Graph of Shear Stress against Normal Stress

From the graph, the interception of the straight line does not pass through zero. Therefore, the cohesion. $c = 2.64 \text{ kN/m}^2$. Besides, the angle of friction is obtained from the plot. From the calculation, the angle of friction of the soil sample is 75.31° . Table 12 shows the results of normal stress and shear stress of soil with root after 30 days for slope gradient 50° .

Table 12: Result of Normal Stress and Shear Stress of soil with root after 30 days for slope gradient 50°

Normal Stress (kN/m^2)	Shear Stress (kN/m^2)
2.78	21.27
5.56	23.21
8.533	27.71

In order to obtain the angle of friction, a graph of shear stress against normal stress was plotted. The shear stress was calculated from the results. The data has been recorded and tabulated in Table 12. The graph has been constructed as shown in Figure 10 and the angle of friction for soil with root after 30 days for slope gradient 50° was determined.

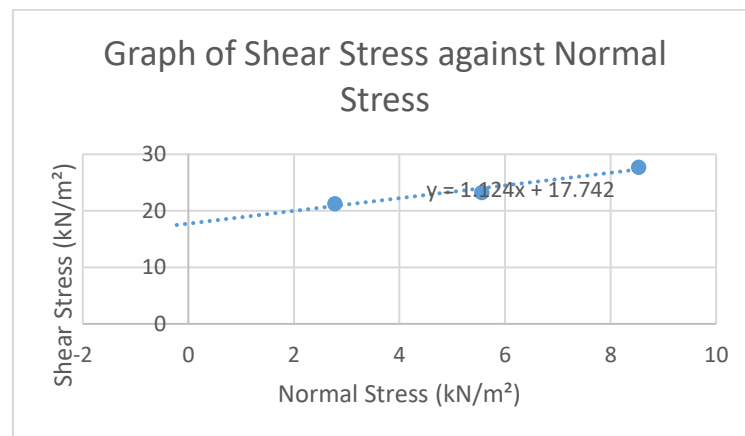


Figure 10: Graph of Shear Stress against Normal Stress

From the graph, the interception of the straight line does not pass through zero. Therefore, the cohesion. $c = 17.74 \text{ kN/m}^2$. Besides, the angle of friction is obtained from the plot. From the calculation, the angle of friction of the soil sample is 48.34° . Table 13 shows the results of normal stress and shear stress of soil with root after 30 days for slope gradient 60° .

Table 13: Result of Normal Stress and Shear Stress of soil with root after 30 days for slope gradient 60°

Normal Stress (kN/m^2)	Shear Stress (kN/m^2)
2.78	21.27
5.56	24.06
8.533	27.22

In order to obtain the angle of friction, a graph of shear stress against normal stress was plotted. The shear stress was calculated from the results. The data has been recorded and tabulated in Table 13. The graph has been constructed as shown in Figure 11 and the angle of friction for soil with root after 30 days for slope gradient 60° was determined.

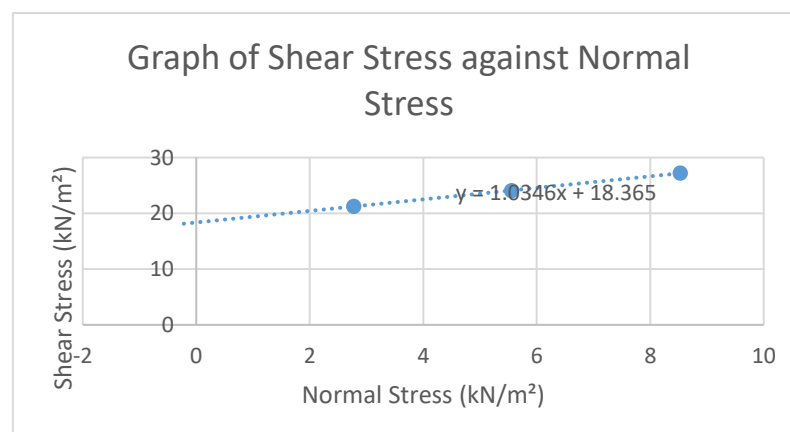


Figure 11: Graph of Shear Stress against Normal Stress.

From the graph, the interception of the straight line does not pass through zero. Therefore, the cohesion. $c = 18.37 \text{ kN/m}^2$. Besides, the angle of friction is obtained from the plot. From the calculation, the angle of friction of the soil sample is 45.97° .

Based on the study, it can be summarized that the strength of soil without the presence of the root of Vetiver grass is weaker compared to that of soil with the support of the Vetiver grass root system. Moreover, based on the angle of slope gradient which are 45°, 50° and 60°, it also effects the strength of soil either with root or without root. The shear strength of soil without root, the highest shear strength for slope gradient 45°, 50° and 60° are 14.95 kN/m², 10.21 kN/m² and 9.48 kN/m². so, slope gradient of 45° has a highest value of shear strength. For 21 days, the highest shear stress of soil and root for slope gradient 45°, 50° are 13.73 kN/m² and 60° is 13.25 kN/m². So, the slope gradient 45° and 50° have the highest shear stress for 21 days. For 30 days, the highest shear stress of soil with root for slope gradient 45°, 50° and 60° is 32.93 kN/m², 27.71 kN/m² and 27.22 kN/m². So, the slope gradient 45° has a highest shear stress for 30 days period. Because vetiver grass has strong roots, it can adhere securely to slopes, increasing the slope's shear resistance. This study indicates that the roots of Vetiver grass could function as a slope-stabilizing reinforcement.

4.3 Tensile Test

As illustrated in Figure 12, the tensile strength towards the root system was shown using a bar graph. In this tensile test, 10 root samples are evaluated for each gradient of slope. The results were determined using two distinct time periods: 21 days and 30 days. The blue bar corresponded to 21 days, whereas the red bar corresponded to 30 days. According to the results that have been analysed, the 30-day-old root has a high tensile strength. For a 45-degree slope gradient, the tensile strength for 30 days is 0.015 kN while for 21 days it is 0.011 kN. For a slope gradient of 50 degrees, the tensile strength for 30 days is 0.020 kN and for 21 days, it is 0.016 kN. For a 60-degree slope gradient, the tensile strength for 30 days is 0.025 kN and for 21 days it is 0.024 kN. According to this study, the Vetiver root with a 60° slope gradient had the maximum tensile strength compared to slope gradients of 50° and 45°.

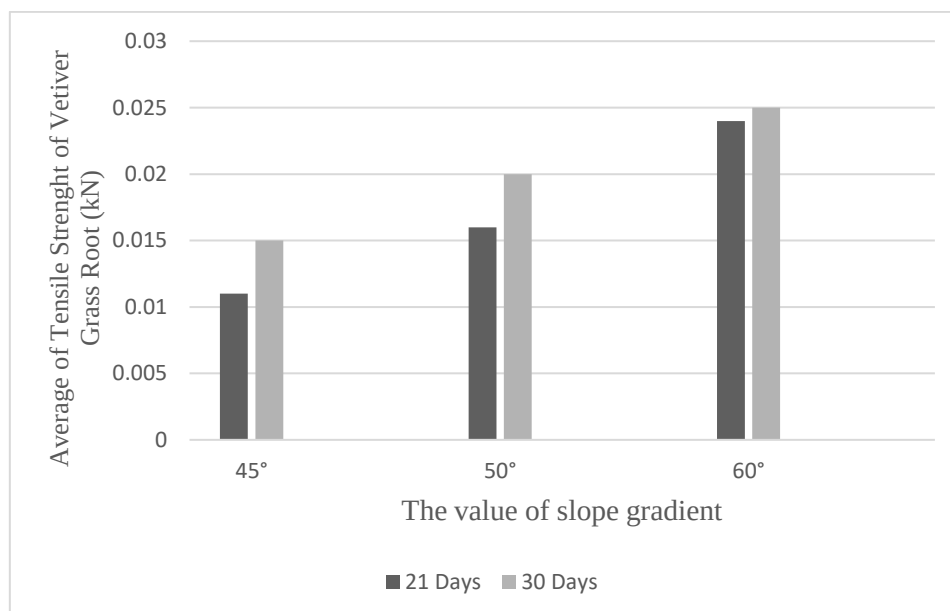


Figure 12: Result of Average of Tensile Strength of Vetiver Grass

5. Conclusion

According to the findings of this study, the presence of Vetiver grass roots greatly affected the soil's strength. Due to the presence of highly erodible residual soils and severe wet weather

Copyright © 2023 Global Academic Excellence (M) Sdn Bhd All Rights Reserved

conditions, Malaysia has made significant strides in the use of Vetiver grass for erosion control and slope stabilization. Based on the collected data, the physical property of the soil was determined. The results of the dry sieve test indicate that the soil type was poorly graded fine to coarse gravel (GP). The information obtained from the cone penetration test utilised in the Atterberg Limit Test was also identified efficiently. At a penetration depth of 20 mm, the liquid limit (LL) was determined to be 20%, and the moisture content ranged from 74.58% to 61.55%. Optimum Moisture Content, OMC demonstrates by Figure 4.2 is 11.3 % striking from the obtained maximum dry density, 1558 kg/m³ respectively. For 21 days, the greatest soil and root shear stress for 45°, 50°, and 60° slope gradients is 13.73 kN/m² and 13.25 kN/m², respectively. Therefore, the 45° and 50° slope gradients have the largest shear stress for 21 days. For 30 days, the greatest shear stress of soil with roots at 45°, 50°, and 60° slope gradients is 32.93 kN/m², 27.71 kN/m², and 27.22 kN/m², respectively. Therefore, the 45° slope gradient has the largest shear stress for a period of 30 days. According to this research, among slope gradients of 50° and 45°, the Vetiver root for 60° slope gradient has the greatest value for tensile strength which are 0.025 kN for 30 days is while for a period of 21 days it is 0.024 kN.

Acknowledgement

This research was funded by a grant from Internal grant Tier 1 (H902). The author would also like to thank the Faculty of Civil Engineering and Built Environment, Universiti Tun Hussein Onn Malaysia for its support.

References

- [1] Bujang, B.K.H.; and Sina, K. (2010). Study of Root Theories in Green Tropical Slope Stability. *Electrical Journal of Geotechnical Engineering*. Vol.15
- [2] Comino, E.; Druetta, A. (2010). The effect of Poaceae roots on the shear strength of soils in the Italian alpine environment. *Soil Till. Res.* 2010, 106, 194–201. <https://doi.org/10.1016/j.still.2009.11.006>
- [3] Coppin, N.J.; and Richards, I.G. (2007). Use of vegetation in civil engineering. London: Construction Industry Research and Information Association, 23-36.
- [4] Graf, F.; Frei, M.; Böll, A. (2009). Effects of vegetation on the angle of internal friction of a moraine. *For Snow Landsc. Res.* 2009, 82, 61–77.
- [5] Hashim, M., Daud, N. M. F. N., Said, Z. M., Nayan, N., Othman, Z., Saleh, Y., Mahat, H. (2018). An Analysis of the Collapse Potential of Slope using the Rom Scale: A Case Study of Sultan Azlan Shah Campus, Sultan Idris Education University, Malaysia. *International Journal of Academic Research in Business and Social Sciences*, 7(6). <https://doi.org/10.6007/ijarbss/v7-i6/3041>
- [6] Islam, M.S. Arifuzzaman & Nasrin. S. (2010). In-situ shear strength of Vetiver Grass rooted soil. *Proc., Bangladesh Geotechnical Conference 2010: Natural Hazards and Countermeasures in Geotechnical Engineering*, 274-279.
- [7] Khalilnejad, A., Faisal Hj. Ali & Osman, N. (2012). Contribution of the Root to Slope Stability. *Geotechnical and Geological Engineering*, 30(2), 277– 288.
- [8] Kosmas, C., Kirkby, M. and Geeson, N. (1999). Manual on: Key indicators of desertification and mapping environmentally sensitive areas to desertification. European Commission, Energy, Environment and Sustainable Development, EUR 18882, 87 p.
- [9] M N Noorasyikin¹ and M Zainab (2016). A Tensile Strength of Bermuda Grass and Vetiver Grass in Terms of Root Reinforcement Ability Toward Soil Slope

- Stabilization. Faculty of Civil Engineering, Universiti Teknologi Mara MALAYSIA
Geotechnical Engineering and Georisk Management, Institute for Infrastructure
Engineering and Sustainable Management, Universiti Teknologi Mara MALAYSIA.
<https://doi.org/10.1088/1757-899X/136/1/012029>
- [10] Mohd Taib, Aizat & Taha, Mohd & Abd Rahman, Norinah & Mat Yazid, Muhamad & Khoiry, Muhamad Azry. (2020). The Effect of Soil-Root Interaction by Vetiver Grass on Slope Stability. *Journal Of Engineering Science and Technology*. Special Issue 3, 46-57
 - [11] Nikhil Kumar, Loveson V. J. and T. N. Singh, Nov 1 – 3. (1998). Change in nutrient status coal overburden dump top material after vegetation – An experimental study. *International Conference on Environment & Agriculture*, Kathmandu, Nepal.
 - [12] Osman, N.; and Barakbah, S.S. (2006). Parameters to predict slope stability. Soil water and root profiles. *Ecological Engineering*, 28(1), 90-95.
<https://doi.org/10.1016/j.ecoleng.2006.04.004>
 - [13] Schwarz, M. Preti, F. Giadrossich, F.; Lehmann, P. Or, D (2010). Quantifying the role of vegetation in slope stability: A case study in Tuscany (Italy). *Ecol. Eng.*, 36, 285–291. <https://doi.org/10.1016/j.ecoleng.2009.06.014>
 - [14] Yoav, W., Amran, E & Katkafi. (2002). *Plant roots: the hidden half*. Book. ISBN: 0824796853. Publisher: Marcel Dekker
 - [15] Yong, W., Qiang, X., & Ying, Z. (2011). Mechanism Study and Numerical Simulation on Vegetation Affecting the Slope Stability. *Electrical Journal of Geotechnical Engineering*. [2011]. Ppr11. 054. Vol 16. Bund G.

THE PERFORMANCE OF VETIVER ROOT GROWTH WITH EM PB AND COCONUT FIBRE – PHYSICAL MODEL

Nurul Izza^{1*}

Nur Najwa²

Noorasyikin Mohammad Noh³

^{1,2,3} Faculty of Civil Engineering and Built Environment, Universiti Tun Hussein Onn Malaysia, 86400 Parit Raja, Batu Pahat, Johor, Malaysia

Abstract: Slope failure can cause serious problems such as infrastructure damage or death. Vegetation is one of the applications used to help slope stability. The problem is slope failure still occurs due to lack of shear strength or tension in the soil. This research aims to overcome the failure with the objective of identifying the morphology and the properties of the vetiver roots, measuring the shear strength and root tensile strength in the presence of EM. PB and comparison between two different planting methods of vetiver grass. Direct shear test and tensile test are conducted after one month and two months to find the tensile and strength of vetiver root that has been planted in PVC pipe. The average growth of vetiver grass plant with soil has the highest value compared to planting with coconut fibre with total of 111.9 cm of growth after 2-month planting compared to with coconut fibre 92.88cm. The result of comparison between two different method of planting shows the value for shear strength after one-month planting for vetiver grass plant with soil and coconut fibre have the highest value with 1.4936 N/mm² compared to vetiver without coconut fibre with the highest value are 1.3671 Nmm². While after 2 months, the highest value is 1.4936 N/mm² with coconut fibre and without coconut fibre are 1.4916 N/mm². Result for tensile after 1-month planting show vetiver root plant without coconut fibre have the highest value of tensile strength with 0.0287 N/mm² while with coconut fibre the highest value is 0.0214 N/mm². Planting in coconut fibre can improve the soil's shear strength contrast to the tensile strength data. By adding EM. PB solution to the soil, agriculture is improved while additionally promoting the development of vetiver grass roots, which can act as reinforcement for slope stabilization.

Keywords: Slope failure, vegetation, root reinforcement, Vetiver grass, slope stabilization

1. Introduction

A slope is an inclined soil mass surface, and they can be created by nature or by humans. Investigations into the slope's shear strength, balance, and ability to withstand movement are particularly crucial to preventing or at the very least minimising slope failure. Slope failure occurs when a slope collapses due to the soil's inadequate ability to support itself, which results from rainfall, earthquakes, and other natural factors. Translational failure, compound failure, wedge failure, and rotational failure are the four types of slope failure that can occur on a site.

Vegetation can be said to a thing that helps in the stability of a slope by making the vegetation reinforce and strengthen the structure of the slope. But the strength will vary according to the character of the plant morphology used as well as the strength of the plant roots themselves to withstand the load. The use of vegetation for slope stabilization is no stranger to some countries

out there and there are studies that have been conducted to see and know the strength of the vegetation on the soil.

This vegetation has many effects on the soil such as tensile strength and shear for example. According to [1] in a study conducted by them in August 2008, "Shear Strength of a soil containing Vegetation Roots, an increase in cohesion to the soil can be seen as the roots of the plant lengthen and this shows that the plant roots respond to cohesion and this depends on the depth of length the root grows". It can be concluded here that the use of vegetation can help in increasing shear strength in soil. In addition, the use of vegetation can also control erosion to the soil by acting as a medium of water absorption. Saturated soil can result in landslides. With the presence of vegetation, it can get rid of the volume of water that does not need to be used for the soil and stabilize the soil condition.

Engineers have long utilised plants when building slopes. Even though putting plants on the slope can assist stabilise it, slope failure still occurs. The absence of shear or tensile strength could be the core of the problem, or it could be that the plant being planted takes a very long time to develop the necessary strength to increase soil stability. Shear stress is a major element in geotechnical studies since we need to know the slope's safety factor value. For the studies, vetiver grass was used as the plant medium. This research is to identify the morphology and the properties of the vetiver roots, measuring the shear strength and root tensile strength in the presence of EM. PB and comparison between two different planting methods of vetiver grass.

2. Literature Review

2.1 Slope failure in Malaysia

According to the study, which is a preliminary investigation into the slope failure at highland in peninsular Malaysia, excessive infiltration brought on by strong rainfall would cause clay particles in the soil structure to swell and cause slope failure [1]. Clay swells and turns porous when it is exposed to water for a long time. In this circumstance, the clay's ability to bind would diminish, weakening the remaining soil structure. A landslide would occur as a result of the collapse of slope stability brought on by the loss of soil strength. Slope failure can be caused by improper slope design, inadequate maintenance, soil type, and gradient slope [2]. A landslide caused by slope failure in Bukit Antarabangsa in December 2008 resulted in 20 more landslides and the deaths of several residents [3].

2.2 Application of vetiver grass

To minimize landslides and erosion, vetiver grass was planted on the banks in the form of terraced plantings, which also helped to catch silt during rainwater runoff. Although vetiver grass has less moisture in its roots than Bermuda grass, which is bad for tensile strength where the value needs to be high, vetiver grass may survive during periods of heavy rain better than Bermuda grass [4]. By transferring shear stress from the soil matrix to tensile inclusions, vetiver roots strengthen the soil. The root matrix increases the soil's shear strength. The root matrix increases the soil's shear strength. The interaction between the soil and roots produces a composite material in which the roots are fibres with a relatively high tensile strength, and here is where vetiver roots play a significant function as a medium to raise the strength of the soil. [5].

In the tropics, the usage of vetiver grass (*Chrysopogon zizanioides*) has been heavily marketed as an economical and environmentally responsible method of slope stabilisation and erosion

prevention. A straightforward technique or practical known as the vetiver system can be employed to aid in soil stabilisation, sediment control, or even water conservation for the soil. On the banks of the land, vetiver grass sown in stages can increase stability and reduce water flow so that it doesn't happen again when it rains. The World Bank first introduced the use of vetiver grass as a slope protection method in India in the 1980s [6].

2.3 Em. Pb Solution for Agriculture Purpose

Strong phototropic bacteria are present in the EM. PB solution. Light is the source of energy for phototropic bacteria, which are widely distributed in nature. harmless to the environment and non-toxic. These microorganisms produce metabolites, which are readily absorbed by plants. These microorganisms operate as nitrogen binders and multipliers of other microorganisms. Use EM; it has been shown that PB with EM1 is more efficient. Benefits of EM. PB include its ability to improve soil and water quality. Aside from that, it will promote healthy plant growth, yield more flowers, and produce fruit that is sweeter. Additionally, the EM. PB solution aids in the breakdown of pollutants and organic waste in water.

2.4 Coconut Fibre for Planting

Coconut fibre or usually known as coconut fibre was a waste product of the coconut industry and was used as a growing medium. Besides, coconut fibre is an environmentally sound peat substitute in soilless growing medium for crop production due to the suitable physical and chemical properties. Coconut fibre might be an alternative, renewable, and reliable growing media for planting because the coconut fibre had a plant growth-promoting bacteria (PGPB) that can enhance plant growth and protect plants from disease and abiotic stresses. Composting of coconut coir dust increased the number of lignin and cellulose degradation bacteria; moreover, it indicated that coconut coir dust is potentially for growing medium with high activity of plant growth promoting bacteria [7].

3. Materials and Methods

The experiments that have been conducted are direct shear stress test and tensile test for root samples. The effectiveness of vetiver grass as well as EM. PB solution was seen through the results of shear test and tensile test

3.1 Determination of Root Properties

Root properties were done, and the root morphology, diameter and height were obtained. To obtain the root, vetiver grass is planted as shown in Figure 1. Equipment or compounds that were used for physical modelling include PVC, soil, vetiver grass, EM, PB chemical and hand shovel. The PVC pipe was cut to a predetermined length of 500mm. Then, soil was inserted into PVC using a hand shovel. The space for the vetiver grass was left. The vetiver grass was placed vertically inside the PVC and added soil until the roots were covered. Sample divided into two categories where EM. PB solution (disturbed) with percentage of 5% and 10% and the other one without solution added (undisturbed). After 1 month and two months of planting, three samples of root were brought to the laboratory to run the test.



Figure 1: Vetiver grass was planted in PVC pipe – Physical Modeling

3.2 Direct Shear Test

The equipment's that were used in this testing are shear box container, base plate with cross groves on top, porous stone, plain grid plate, perforated grid plates, loading pad, weighing, load frame, dial gauge, tampering rod, spatula, rammer, proving ring and soil sample. For this testing, the soil was divided into three parts. The sample was put aside first and the shear test apparatus was set up first. The testing was started after the set up was finished and the readings are recorded for every 20 drops until three constant readings are obtained. The loading started with 2kg of load. The shear box was unlocked, and the procedure was repeated for 4kg and 6kg load. The data was put on the table to calculate the shear stress, normal shear and after that plot the data to obtain cohesion, gradient and lastly shear strength by using formula. The shear strength is equal to the sum of the normal stress (σ) and cohesion (c). The Mohr coulomb equation is shown as below in Eq.1:

$$\text{Shear Strength } (T) = C + \sigma n \tan \theta \quad (1)$$

4. Result and Discussion

The experiment was carried out after the planted vetiver grass had enough planting time according. The study that has been conducted is tensile test using Universal test machine and direct shear test.

4.1 Root Properties Determination

Root properties was done, and the root morphology, diameter and height were obtained shown in Figure 2. The length of grass was measured every once a week and the data were recorded. For diameter and the length of the root, the measurement was taken before testing are run. Table 1, 2, 3, and Table 4 shows the average length growth of vetiver grass. The highest vetiver growth is at the average of 70.2 cm. Vetiver grass growth increased by 30 to 60 percent over the course of a month for every type of sample. For controlled sample with zero percent of EM. PB solution, the average grass growth is 68.7 cm. For sample with five percent of EM. PB solution, the average grass growth is 69.3 cm. And with ten percent solution, the total average grass growth is 70.2 cm. For the first week, all vetiver grass grew over 15 cm with the average of 21.5 cm.

For Table 2, The highest average of growth is at vetiver grass with the presence ten percent of EM. PB solution with value of 111.9 cm. Average growth for controlled sample is 100cm per

month. For uncontrolled sample, 109.5 cm for five percent and 111.9 cm for ten percent of solution. The vetiver grass growth is higher when the EM. PB solution are added per week. From the result, it shows that the presence of EM. PB solution helps in vetiver growth.



Figure 2: Determination of Root Morphology

Table 1: Average length growth of vetiver grass for 1 month

EM. PB Solution	Week	Average Length growth (cm)
0	1	15.6
	2	28.3
	3	46.6
	4	68.7
5	1	24.6
	2	41.0
	3	61.7
	4	69.3
10	1	24.2
	2	51.7
	3	59.1
	4	70.2

Table 2: Average length growth of vetiver grass for 2 months

EM. PB Solution (%)	Week	Average Length growth (cm)	Week	Average Length growth (cm)
0	1	16.5	5	79.4
	2	30.0	6	85.7
	3	49.4	7	92.6
	4	72.8	8	100.0
5	1	26.1	5	80.1
	2	43.5	6	88.9
	3	65.4	7	98.7
	4	73.5	8	109.5
10	1	25.7	5	79.6
	2	54.8	6	89.2
	3	62.6	7	99.9
	4	74.4	8	111.9

Table 3 shows the average growth of the coconut coir vetiver grass plant after one month. The average grass growth for a controlled sample with 0% EM. PB solution is 58.40 cm. The average grass growth for the sample containing 5% EM. PB solution is 58.91 cm. Additionally, with a 10% solution, the overall average grass growth is 59.67 cm. Every vetiver grass grew more than 10 cm within the first week.

Table 3: Average length growth of vetiver grass for 1-month plant with coconut coir

EM. PB Solution	Week	Average Length growth (cm)
0	1	13.26
	2	24.06
	3	39.61
	4	58.40
5	1	20.91
	2	34.85
	3	52.45
	4	58.91
10	1	20.57
	2	43.95
	3	50.24
	4	59.67

Table 4: Average length growth of vetiver grass for 2-month plant with coconut coir

EM. PB Solution (%)	Week	Average Length growth (cm)	Week	Average Length growth (cm)
0	1	13.695	5	65.90
	2	24.9	6	71.13
	3	41.002	7	76.86
	4	60.424	8	83.00
5	1	21.663	5	66.48
	2	36.105	6	73.79
	3	54.282	7	81.92
	4	61.005	8	90.89
10	1	21.331	5	66.07
	2	45.484	6	74.04
	3	51.958	7	82.92
	4	61.752	8	92.88

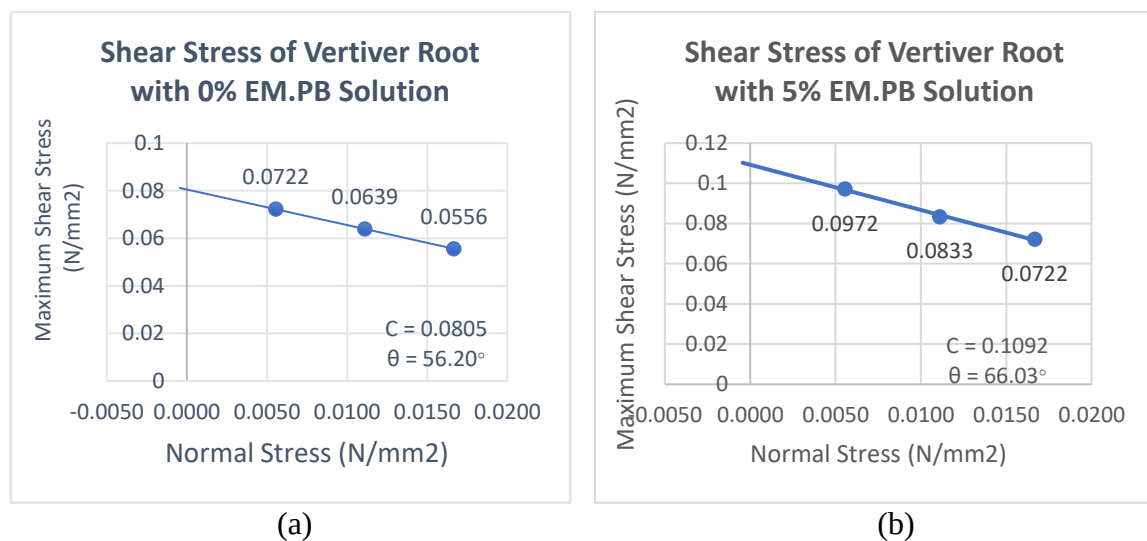
Following a two-month planting period with coconut coir, Table 4 displays the average growth of vetiver grass. The vetiver grass, which contains ten percent EM. PB solution and has a value of 92.88 cm, has the highest average growth. After two months, the controlled sample's average growth is 83.00 cm. For an uncontrolled sample, the average height of vetiver grass was 90.89 cm for a five percent solution and 92.88 cm for a ten percent solution. When the EM. PB

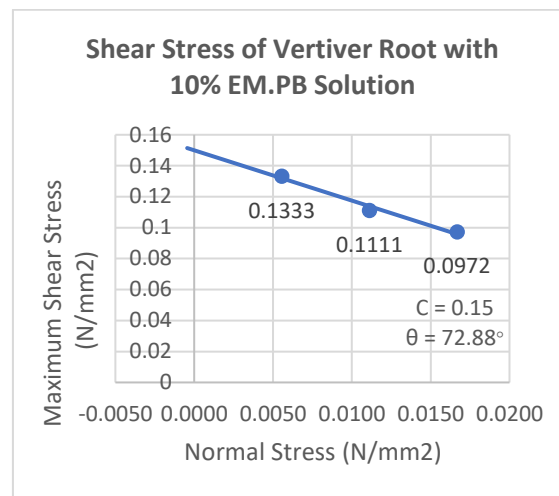
solution is added weekly, the vetiver grass grows more quickly. The outcome indicates that the presence of EM. PB solution promotes the growth of vetiver.

4.2 Direct Shear Test

Direct shear test was performed with three (3) different conditions where each sample had a different value of additives and value of loading. Figure 3 and 4 show the value of shear strength for the one-month planting and Figure 5 and 6 show shear stress and shear strength value of sample after two-month planting with only soil and addition of EM. PB solution. Each graph represents the result with a different amount of EM. PB solution added. Data from the graph are calculated to obtain shear strength value.

The cohesion for graph (a) is 0.0805 kPa. It can be seen through the three graphs above; the highest amount of shear strength is at the top of the soil. For a sample with zero percent of EM. PB solution, the highest value of shear strength is at the top part of the soil with value of 1.019 N/mm², while the maximum shear stress for the soil is 0.0722 N/mm². From graph (b), the cohesion for the soil sample is 0.1092 kPa and the highest value for maximum shear stress is 0.0972 N/mm². The highest value for shear strength is 1.2119 N/mm². From the graph (c), the value cohesion is obtained which is 0.15 kPa. This value is used to find the shear strength by using formulas. For a sample with ten percent of EM. PB solution, the highest shear strength is also at top of the soil with value of 1.3671 N/mm². The highest maximum shear stress is 0.1333 N/mm².

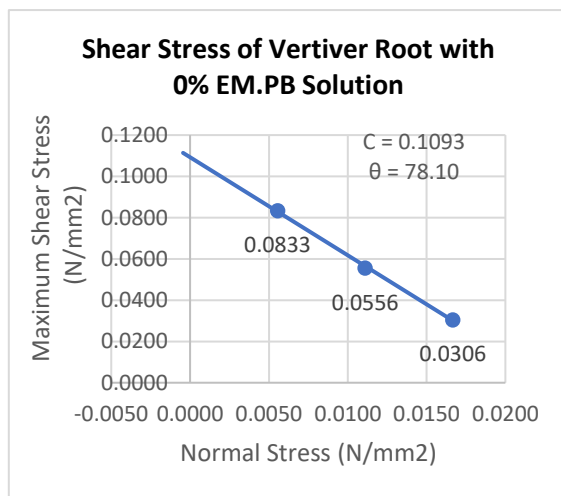




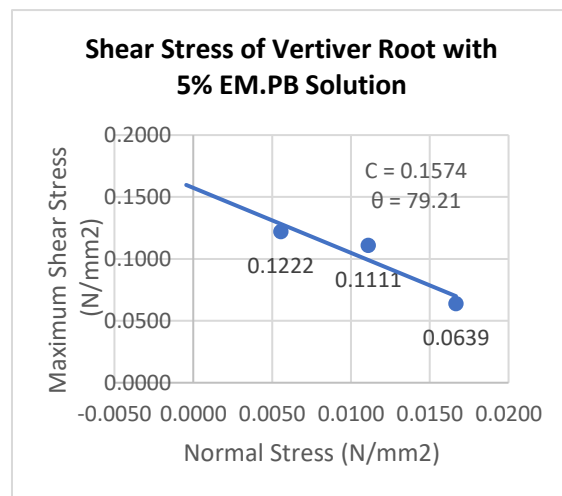
(c)

Figure 3: Shear Strength Analysis After 1-Month Planting

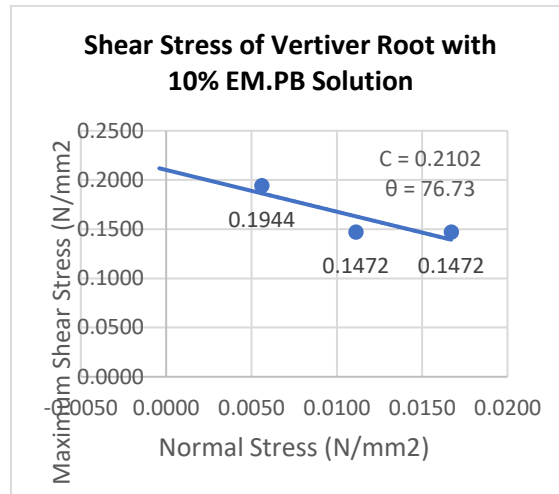
From graph (b), cohesiveness for the soil sample is 0.1574, and the greatest value for maximum shear stress is 0.1222 N/mm² and the maximum shear strength value is 1.4802 N/mm². The value of cohesiveness is 0.2102 kPa, as shown by the graph (c) with 0.1944 N/mm² for the highest maximum shear stress. With a value of 0.6399 N/mm², the lowest shear strength is found at the bottom of the soil. The maximum shear stress for bottom and middle soils is 0.1472 N/mm².



(a)



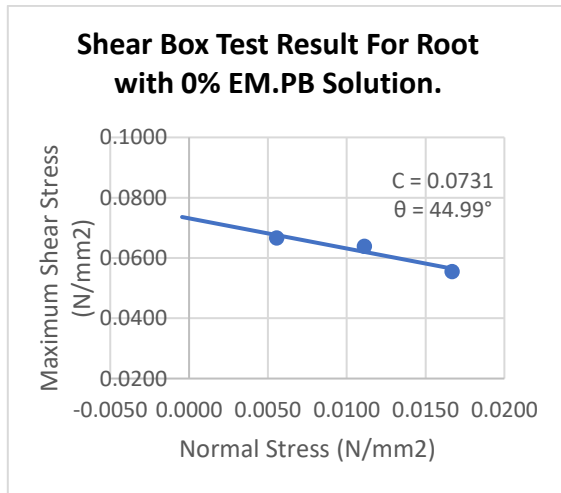
(b)



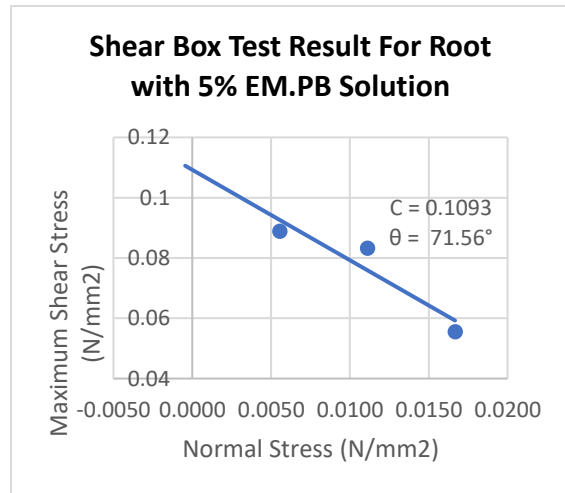
(c)

Figure 4: Shear Strength Analysis After 2-Month Planting.

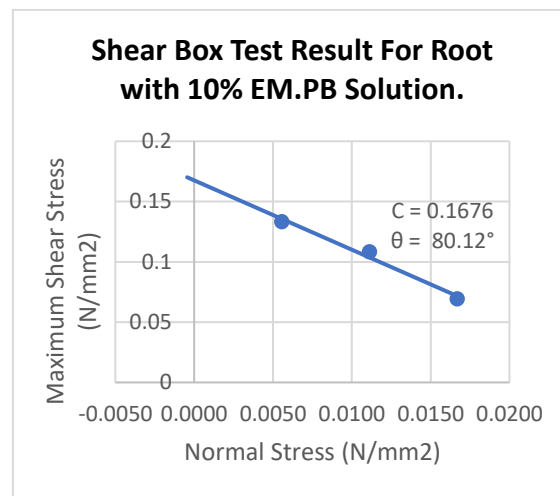
Figure 5 below shows the graph of shear test for vetiver root that is a plant with soil and coconut fibre and addition of EM. PB solution after 1 month of planting. The cohesion for graph (a) is 0.0731 kPa. It can be seen through the three graphs above; the highest amount of shear strength is at the top of the soil. For a sample with zero percent of EM. PB solution, the highest value of maximum shear stress for the soil is 0.0667 N/mm². From graph (b), the cohesion for the soil sample is 0.1093 kPa and the highest value for maximum shear stress is 0.0889 N/mm². From the graph (c), the value cohesion is obtained which are 0.1676 kPa with 0.13333 N/mm² value of maximum shear stress.



(a)



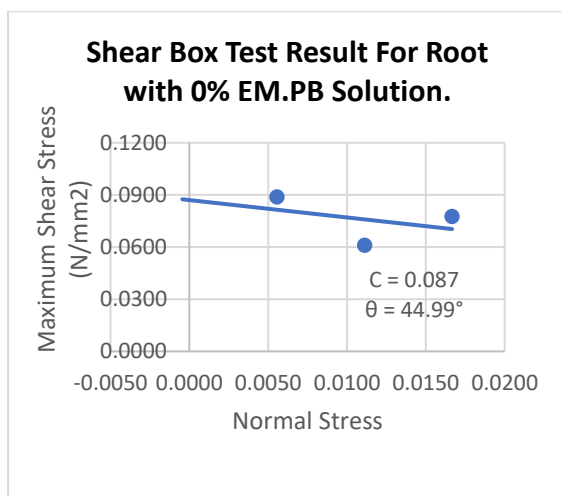
(b)



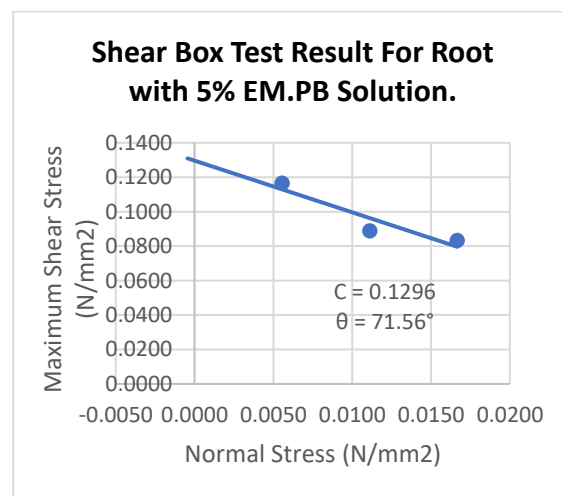
(c)

Figure 5: Shear Strength Analysis After 1-Month Planting

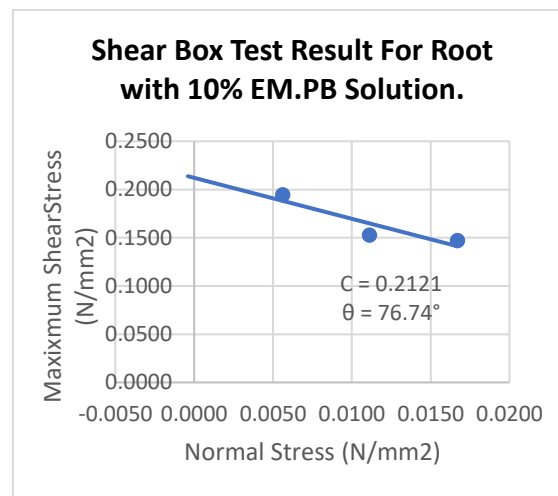
Figure 6 shows the result of shear stress for vetiver root plant with the presence of coconut fibre and soil after 2 months. Value of cohesion for graph (a) is 0.087 with 0.0889 N/mm² of maximum shear stress. From graph (b), cohesiveness for the soil sample is 0.1296, and the maximum shear strength value is 0.1167 N/mm². The value of cohesiveness is 0.2121 kPa, as shown by the graph (c) with 0.1944 N/mm² for the highest maximum shear stress.



(a)



(b)



(c)

Figure 6: Shear Strength Analysis After 2-Month Planting

Based on the result, the value of shear strength is higher when the additive solution is added. This could be seen from the different findings between zero percent of EM. PB solution and ten percent of EM. PB solution added during planting.

Table 5 and table 6 shows the comparison of data between vetiver grass plant with soil and soil that is mixed with coconut fibre. From the table, vetiver grass plants with zero percent of EM. PB solution, the highest shear strength is 1.019 N/mm² which is a vetiver grass plant with only soil. For vetiver grass with five percent of EM. PB solution, the highest shear strength is 1.3043 N/mm² and for ten percent solution, highest value is 1.5057 N/mm². Both have the highest value of shear strength for five and ten percent of EM. PB solutions are vetiver grass plants with the presence of coconut fibre and soil.

Table 5: Value Of Shear Strength Result After One-Month Planting

EM. PB Solution (%)	Load (kg)	Soil Part	Plant with Soil and coconut fibre		Plant with soil	
			Max Shear Stress	Shear Strength, N/mm ² (T)	Max Shear Stress	Shear Strength, N/mm ² (T)
0	2	Top	0.0667	0.8245	0.0722	1.019
	4	Middle	0.0639	0.5725	0.0639	0.743
	6	Bottom	0.0556	0.3251	0.0556	0.3952
5	2	Top	0.0889	1.3043	0.0972	1.2119
	4	Middle	0.0833	0.9036	0.0833	0.8421
	6	Bottom	0.0556	0.51	0.0722	0.479
10	2	Top	0.1333	1.5057	0.1333	1.3671
	4	Middle	0.1083	1.057	0.1111	0.959
	6	Bottom	0.0694	0.6163	0.0972	0.5581

Table 6 shows the result of vetiver grass after 2-month planting. From the table, vetiver grass plants with zero percent of EM. PB solution, the highest shear strength is 1.4136 N/mm² which is vetiver grass plant with soil and coconut fibre. For vetiver grass with five percent of EM. PB

solution, the highest shear strength is 1.4802 N/mm² and for ten percent solution, highest value is 1.4936 N/mm².

Table 6: Value of Shear Strength Result After 2-Month Planting

EM. PB Solution (%)	Load (kg)	Soil Part	Plant with soil and coconut fibre		Plant with soil	
			Max Shear Stress (N/mm ²)	Shear Strength, N/mm ² (T)	Max Shear Stress (N/mm ²)	Shear Strength, N/mm ² (T)
0	2	Top	0.0889	0.8384	0.0833	1.4136
	4	Middle	0.0611	0.5864	0.0556	0.9762
	6	Bottom	0.0778	0.339	0.0306	0.5467
5	2	Top	0.1167	1.3246	0.0639	1.4802
	4	Middle	0.0889	0.9239	0.1222	1.0366
	6	Bottom	0.0833	0.5303	0.111	0.601
10	2	Top	0.1944	1.4936	0.1944	1.4916
	4	Middle	0.1528	1.0639	0.1472	1.0619
	6	Bottom	0.1472	0.6418	0.1472	0.6399

According to the findings, adding the additive solution increases the shear strength value. This was evident from the disparate results between the addition of 0% EM. PB solution and 10% EM. PB solution during planting. After increasing the planting time, the vetiver grass plant with the presence of coconut fibre and soil shows higher value of shear strength compared to only soil. Planting with coconut fibre does help in increasing the shear strength of the soil

5. Conclusion

The results of the shear stress test indicate that the presence of the EM. PB solution contributes to enhancing the shear value. According to the findings, the bottom part of the root has lower shear stress resistance compared to the top and middle bottom. The shear strength decreases as the root section gets smaller. The root which has no EMPB solution present had the highest tensile strength result, as per the results of the test. This demonstrates that the tensile results are unaffected by the presence of the EM. PB solution in the root. As for the vetiver grass plant with addition of coconut fibre, after extending the planting period, the vetiver grass plant with soil and coconut fibre exhibits more shear strength than one with just soil. The soil's shear strength can be increased by planting in coconut fibre. Different from the results of the tensile strength where the vetiver grass plant soil shows higher tensile value compared to vetiver grass plant with presence of coconut fibre. Planting with coconut fibre does not help in increasing the tensile strength for the vetiver root. The application of EM. PB solution to the soil can be seen as an additional chemical that improves agriculture while also promoting the growth of vetiver grass roots, which can work as reinforcement for slope stabilisation.

Acknowledgement

This research was funded by a grant from Internal grant Tier 1 (H902). The author would also like to thank the Faculty of Civil Engineering and Built Environment, Universiti Tun Hussein Onn Malaysia for its support.

References

- [1] Faizal H.A, and Normaniza O. Soil and Foundation 4th Edition. Review of Shear Strength of a Soil Containing Vegetation Roots. *Japanese Geotechnical Society*, vol. 48, pp. 587–596, 2008.
- [2] Gasim, M.B,Sahibin, A.R and Toriman,M.I. (2011).Slope Failure at Bukit Antarabangsa, Ampang, Selangor, and Its Relationship to Physical Soil Properties. *The Malaysian Journal of Analytical Sciences*, vol. 15, pp. 138–149, 2011.
- [3] Mukhlisin, M., Matian,S.J., Taha,M.R and Ahlan,M.J (2015). Analysis of Rainfall Effect to Slope Stability in Ulu Klang, Malaysia. *Jurnal Teknologi*, vol 72, no. 3.<https://doi: 10.11113/jt. v72.4005>.
- [4] Noorasyikin M. N and Zainab.M. (2016). A Tensile Strength of Bermuda Grass and Vetiver Grass in Terms of Root Reinforcement Ability Toward Soil Slope Stabilization. IOP Conference Series: Materials Science and Engineering, vol. 136, pp. 012029. <https://doi: 10.1088/1757-899x/136/1/012029, 2016>.
- [5] Osano,S.N & Mwea, S.K (2011), Root tensile strength of 3 typical plant species and their contribution to soil shear strength; a case study: Sasumua Backslope, Nyandarua District, Kenya. *Journal of Civil Engineering Research and Practice*. Vol.8, No.1,pp 57-73. <https://doi:10.4314/jcerp.v8i1.69525>
- [6] Osano,S.N, Mwea,S.K & Gichaga, F.J. (2012). Pull-Out Resistance of 3 Different Plant Species and Their Application in Slope Stabilization Works. *Journal of Engineering Indian Centre for Advanced Scientific & Technological Research (Icator)*. Vol.5 (1)
- [7] Osman, N.; and Barakbah, S.S. (2006). Parameters to predict slope stability. Soil water and root profiles. *Ecological Engineering*, 28(1), 90-95. <https://doi.org/10.1016/j.ecoleng.2006.04.004>
- [8] Preti,F & Giadrossich,F (2009). Root reinforcement and slope bioengineering stabililzation by Spanish Broom (*Spartium junceum* L.). *Hydro.Earth Syst. Sci. J*. Vol.13. pp 1713 – 1726. <https://doi.org/10.5194/hess-13-1713-2009>
- [9] Rahardjo,H., Leong,E.C & Rezaur,R.B (2002). Instrumented slopes for the study of rainfall induced slope failures. National Seminar Slope 2002. 27 April 2002. Proceedings of the National Seminar. <https://doi.org/10.1016/j.sandf.2020.11.006>
- [10] Sasan,M., Bujang, B.K.H., & Vahed, G. (2009). Evaluation on Root Theories and Root Strength Properties in Slope Stability. *European Journal of Scientific Research*.Vol.30 No.4.594-607. ISSN 1450-216X
- [11] Stephan, H., Fabricio,J.S & Hans, P.R. (2013). A pull out test to compare two riparian species, *Phyllanthus sellowianus* and *Sebastiania schottiana* in terms of root anchorage ability. *Geophysical Research Abstracts*. Vol.15, EGU2013-
- [12] Schwarz, M. Preti, F. Giadrossich, F.; Lehmann, P. Or, D (2010). Quantifying the role of vegetation in slope stability: A case study in Tuscany (Italy). *Ecol. Eng.*, 36, 285–291. <https://doi.org/10.1016/j.ecoleng.2009.06.014>
- [13] T K, D. (2020). Study of Plant Growth Promoting Bacteria from Coconut Coir Dust. *IOP Conference Series: Earth and Environmental Science*. <https://doi: 10.1088/1755-1315/439/1/012037>
- [14] Yong,W., Qiang,X., & Ying, Z. (2011). Mechanism Study and Numerical Simulation on Vegetation Affecting the Slope Stability. *Electrical Journal of Geotechnical Engineering*. [2011]. Ppr11. 054.Vol 16. Bund G.

- [15] Zabidi, M. D., Hadi, B. A., K.R., A., Keria, R., & Derahman, A. (2021). A Preliminary Study on the Slope Failure at Highlands in Peninsular Malaysia. *International Journal of Academic Research in Business and Social Sciences*, 11(4), 585-594

COMPRESSIVE SENSING PATH FOR OPTIMAL DATA TRANSMISSION IN UNDERWATER ACOUSTIC SENSOR NETWORK

Gunasekar Thangarasu ^{1*}
Kesava Rao Alla ²

¹Department of Professional, Industry Driven Education, MAHSA University, Saujana Putra, Malaysia

²Chancellery, MAHSA University, Saujana Putra, Malaysia

Abstract: *The research team has developed several methods to minimize the power consumption of underwater sensor networks, including clustering, compressive sensing, and the use of a block diagonal matrix. By combining the clustering process with either a block diagonal or compressive sensing matrix, we can significantly reduce the energy required for sending and receiving sensor data. This results in faster transmission and reception of data. To transfer the compressed sensing measurement data from the cluster head to the base station as quickly as possible, a routing strategy is devised based on the shortest possible path. Additionally, arbitrary directions are used to communicate the results of compressive sensing based on clusters. Overall, these techniques offer an efficient way to optimize energy usage and improve the performance of underwater sensor networks.*

Keywords: *Compressive Sensing, Optimal path, Routing, Underwater Sensor Network*

1. Introduction

The use of underwater sensor networks (UASNs), has grown increasingly prevalent in a variety of settings, including those that are military as well as those that are civilian. The situation is analysed after the sensors, which are generally submerged, have been observed following their random distribution around the area that is within the range of their detection. In spite of the fact that they function most of the time without the need for routine maintenance or the utilisation of renewable energy sources, they are still able to carry out the duty that has been assigned to them [1].

Low-cost, low-power, and tiny devices are an absolute requirement for the operation of underwater sensor networks and their interaction with one another. During the entirety of the process of data collecting, preserving power in such networks has always been a key challenge. This is since it has a direct impact on the ability of the network to function over the course of time.

According to several studies that highlight the significance of this finding, keeping better track of data in underwater sensor networks consumes less power than was previously anticipated (UASNs). The process through which network sensors are placed in sensing hotspots is entirely arbitrary. The majority of the time, they are put to service in hazardous environments that do not have access to sources of electrical energy, nor do they receive routine maintenance. Because of the severe restrictions placed on the available energy, the network connection is forced to rely on extremely small and inexpensive equipment [2].

Compressive sensing (CS), is a novel method of data processing that was developed specifically for signals that are sparse or compressible. This approach offers to reconstruct all sensor readings from N nodes using just $M = O(k \log N/k)$. CS measurements gathered from the sensing area and communicated to the sink or base station because the sensor data in UASNs is spatially correlated (BS). This approach achieves this objective because the sensory input in UASNs is spatially associated with one another [3].

Sensory data is typically encoded as a k -sparse signal using the canonical basis or one of several alternative sparsifying bases, such as the discrete cosine transform (DCT) or the wavelet transform. This encoding method is utilised so that the data can be processed in a more efficient manner. Each CS measurement can be gathered either from all the sensing nodes in the network or from selected random nodes, and the decision of which approach to use is determined by the data collecting methods that are utilised in the network that employs CS. In recent years, a vast number of research projects about computer science and the collecting of data in UASNs have been carried out. These projects have been carried out all over the world.

Several routing methods, the most notable of which are gossip-based, random-walk, tree-based, and cluster-based systems, contribute, at least in part, to the ease with which sensor values can be transmitted to the BS. At the BS, we gather the CS data in the form of $Y_{M \times 1} = \Phi_{M \times N} \times N \times 1$, where m is the measurement matrix (which is also known as the projection matrix), and X represents all the unknown values from all of the sensors. The data are collected in this format so that they can be analysed by the projection matrix. The underlying routing mechanisms are responsible for the variety in the final measurement matrices, which can be either sparse or rich in Gaussian coefficients. This variation is caused by the fact that the routing processes are buried under the surface.

In RW routing that uses CS, it is necessary to send out a predetermined number of RWs of a predetermined duration to collect sensor readings from randomly selected nodes. The length of these RWs is also predetermined. This is essential for UASNs since all sensors are dispersed at random over the sensing area in such networks. To CS recovery, the data gathered by each of the node sensors are consolidated into a single measurement of the CS before being sent to the BS. This is done before the data can be used to recover the CS. After then, this measurement is transmitted. During the process of CS recovery, several sensor nodes use a wide variety of distinct measurement matrices to collect data. Although it has been shown that random sparse measurement matrices are capable of just as much success as complete Gaussian ones in CS recovery methods, the RW routing that we recommend makes use of a sparse binary matrix.

The measurement matrix will have one binary row that won't be used at all. This is because nodes only need to collect sensory input from a subset of the network rather than the complete thing. Over the time span of M iterations, each node gathers a sampling of data at random, which is represented by the equation $Y_{M \times 1} = \Phi_{M \times N} \times N \times 1$. After that, these data are transmitted to the BS so that they can be incorporated into the CS recovery procedure.

The measurements of sensor data almost always have some form of spatial relationship, which causes the data to be naturally sparse on a level playing field. For the application of compressive sensing, it is helpful to use data that is already naturally sparse because this enables less data to be collected at a lower compression rate, which in turn results in less space being taken up by the compressed data. In summary, compressive sensing offers a fresh

approach to the age-old problem of how to reproduce complete sensor readings using only the observations that are allowed by restricted compressive sensing. Compressive sensing offers a wide range of benefits, including this one. There is the potential for a considerable reduction in the total amount of power that is consumed by sensor networks.

2. Related Works

The sparse random projections as the basis for a method of cooperative data sensing and compression. Because of this strategy, the sensors, which were already cooperating to get the intended outcomes, were not required to exert any more effort. By taking advantage of the spatial correlation of the data, the method can not only deliver great reconstruction accuracy for the detected field but also experience a significantly lower communication traffic burden when compared to conventional sampling algorithms that are used in underwater sensor networks [4]. This is possible because the method can take advantage of the spatial correlation of the data. This is conceivable since the method can make use of the spatial correlation that is present in the data. In the trials, the reconstruction accuracy and energy efficiency under different models of traffic movement are tested using data sets obtained from actual metropolitan environments in the real world [5-6].

The fundamental goal of the adaptive CS-based sample scheduling method (ACS) for UASNs, which was created, is to obtain good sensing quality while yet retaining a low sample rate. to arrive at the optimal sample rate for each sampling window, the ACS considers not only the minimum required sample rate that is forecasted, but also the quality of the data that has been sensed. The implementation of ACS in cognitive sensor networks grants the networks the capability to carry out a variety of tasks, including the monitoring of their surroundings and the sensing of the spectrum [7-8].

Compressed sensing as a more methodical answer to the problem. This suggestion comes highly recommended. It is possible to gradually restore the original data while still preserving the benefits of the NC-based approach, and the measurement and recovery strategy that we advocate is what makes all this possible [9-10].

The combination of CS and RW to find ways to mitigate the negative effects that human activity has on the natural environment. It takes fewer CS measurements than the total number of sensor nodes in the network to reconstruct all the sensory data at the BS. A RW route with a duration that has been defined in advance is utilised so that each CS measurement can be obtained. The base station (BS) can gather random CS measurements for the purpose of CS recovery in one of two ways: either directly or through relaying through intermediary nodes. Both methods are described further below. to ensure that the networks use the least amount of energy possible, researchers are examining the best potential compromise that can be made between the sensor transmission range and the length of RWs. This is done to ensure that the networks can function properly [11-12].

Using an asymmetric semi-homomorphic encryption approach and a sparse compressive matrix proposed a method for the secure collection of data that is based on compressive sensing. This method makes use of compressive matrices that are sparse. This was done to ensure the security of the information that was being collected (SeDC). To be more explicit, the asymmetric technique makes it easier to disseminate and handle secret keys [13-14].

The use of homomorphic encryption not only makes it feasible for load balancing to take place within a network, but it also makes the network more secure. Since it enables aggregation within the network in the cypher domain, this is made possible [10]. The sparse measurement matrix can reduce the amount of money spent on processing and communication, hence mitigating the increase in costs that are the result of homomorphic encryption. In addition, we offer a model for collaborative recovery to further enhance the level of precision achieved during the recovery process.

3. Proposed Method

This clustering and creation of compressive sensing measurements utilising block diagonal metrics is made possible through the utilisation of this compressive sensing technique, which then transmits these data to the base station either directly or over a series of hops. Because the compressive sensing measures can be distinct from one another depending on the sensors that are utilised in the network, it is essential for each cluster leader to have their own unique metric requirements.

In accordance with the principles underlying the concept of sensor networks, there are N sensors spread out over an $L \times L$ -sized region. It is necessary for the cluster head to receive the data from the sensor that is not physically positioned within the cluster head since that information is dependent on the distance that is being measured. The cluster heads are the ones in charge of the generation of compressed sensing data, which is subsequently delivered straight to the base station after it has been processed. The location of the base station can be moved to anywhere inside or outside of the sensing region at the user discretion.

Although the non-cluster heads are the ones who are tasked with determining how much the communication will cost, this is since the cluster heads are responsible for the production of the compressive sensing. to generate compressed sensing measures for usage within the cluster, the data from many sensors are merged. The command-and-control centre receives an immediate notification of the readings. The following equation can be used to calculate the total amount of power consumed [15]:

$$P_{total} = (P_{BS} + P_{intra-cluster})$$

where,

$P_{intra-cluster}$ - intra-cluster consumption and

P_{BS} - power consumption.

It is something that can be noted that the $P_{intra-cluster}$ plays a less significant function when the total number of clusters in the system decreases. This is something that can be seen. $P_{intra-cluster}$ is computed with the assistance of the PSO model to reduce the amount of quality that is lost during the process of sending compressive sensing measurements from cluster heads to intermediate cluster heads and then to the base stations.

$$P_{intra-cluster} = \left(\frac{N}{N_c} - 1 \right) \frac{L^2}{2\pi}$$

This is done because a large amount of data is transferred during the process of sending compressive sensing measurements from cluster heads to intermediate cluster heads and then to the base stations. Large clusters often make use of a technique known as inter-cluster hop

routing, which simply means hopping across clusters. Compressive sensing methods require less energy to transmit data as the number of cluster hops (CH) in each network rises. This is made possible by inter-cluster-hop routing.

After the clusters have been created with the assistance of the k-means algorithm, the focus of the inquiry that has been recommended will be on determining the path that is the most time and resource effective to travel from the hub to the base node. An algorithm known as PSO is utilised to successfully complete this task. To establishing a connection between the various cluster nodes and the roots of the base station, this approach makes use of a tree-like structure. It is a widely held belief that all of the nodes that comprise a cluster have the same transmission range (R), and that nodes that are able to communicate with one another and share data with one another are able to do so as long as they are within that range.

The newly created cluster heads are evaluated to make a decision regarding the transmission range that will be utilised. This makes use of a graph that is built on undirected geometry to connect each of the cluster unique nodes with one another in an efficient manner. $G = (V, E)$, where V is the number of cluster heads and E is the set of edges by which they can communicate with one another. The number of cluster heads is denoted by V, and E is the set of edges. The number of cluster heads is represented by the letter V, while the set of edges is marked by the letter E. Both the communication links that exist between cluster heads and the cluster heads themselves are subject to change depending on the circumstances.

Both the constrained variable, also known as the trip time t_{ij} , and the total cost c_{ij} are instances of components of a connection that have positive numerical values. The trip time t_{ij} is another name for the restricted variable. Integer programming is a method that is used for modelling the problem of determining the shortest route between the home station and the cluster leader while adhering to a set of constraints. This is accomplished by modelling the problem as a mathematical equation.

The PSO method can be used to figure out the route that will transfer data from sensors to nodes that are a part of a cluster in the most time and energy-effective manner. Creating a graph is the method used to accomplish this. All the nodes that comprise a cluster have the ability to communicate with one another and exchange information about the distance that separates each node from the base station thanks to the heads that are attached to each node. At first, the cluster will concentrate its activity inside the transmission range of the base station, which will be closer to it. The first thing that has to be done to figure out the maximum number of hops that can be used to transmit the data is going to be this. The value of Hops is initially set to 1, and during each iteration, the current state is communicated to any neighbouring non-cluster heads that are located in the vicinity of the area.

The purpose of developing the PSO model is to find out which types of undirected sensor networks have the quickest routes to reach their final objectives. In this instance, the selection of the shortest path problem is the primary concentration, although the shortest path tree is of just a moderate degree of significance. If there are N nodes in the sensor network, then the path that leads to the sink with the shortest distance will have precisely one node in it. So, the procedure must continue to iterate until the shortest path tree has been built, connecting all of the nodes other than the final one (N-1). Using the PSO technique, it is possible to find a

solution to the other challenge, which entails constructing the shortest path tree for a dynamic graph.

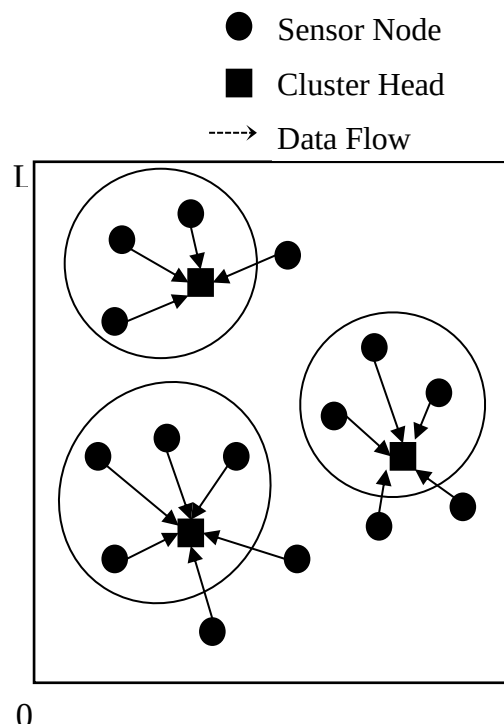


Figure 1: Cluster-Compressive Sensing

Figure 1 illustrates how the suggested procedure is segmented into three distinct stages that can be followed in any order.

- While working with digraphs, the challenge of finding the path with the shortest distance between two points is tackled in the first phase, which makes use of a generalised PSO model. During this stage, any additional issues that may have sprung up are also addressed and resolved.
- In the second step, we make some modifications to the generalised PSO model so that it can manage the complications of optimising a network that has many sinks. This allows the model to produce more accurate results.
- In the third stage, the newly acquired information is used in the process of reconstructing the shortest path tree. Modifications to the link weight provide the fundamental basis for the flow of the link network to be redirected in the desired direction.

4. Results and Discussions

In this section of the essay, we will present a comprehensive study of the method that has been proposed. This method will be evaluated by putting it through its paces using a variety of different databases. This considers both the signals that have not been sorted and those that have been sorted, we are able to arrive at the conclusions that we are looking for. The data collecting process is handled by the sensors that are a member of the square sensing network. Not only can the availability of such data at varying k values influence the reconstruction error that takes place at the base station, but it can also influence the cost that is incurred when

sending data from the cluster head to the base station. Both effects can be caused by the availability of such data.

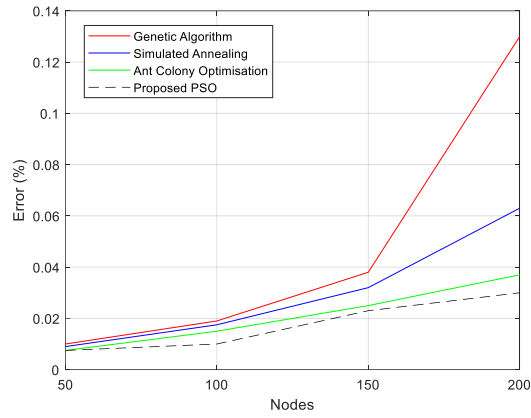


Figure 2: Reconstruction Error

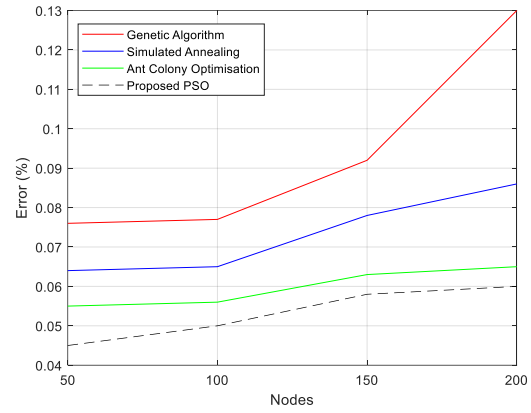


Figure 3: Normalized Reconstruction Error

As in Figure 2, the reconstruction error at the base station has a propensity to grow either when the total number of clusters in the data set grows or when the total number of coefficients that are communicated to the base station decreases. Both scenarios are shown to have a negative impact on the accuracy of the reconstruction.

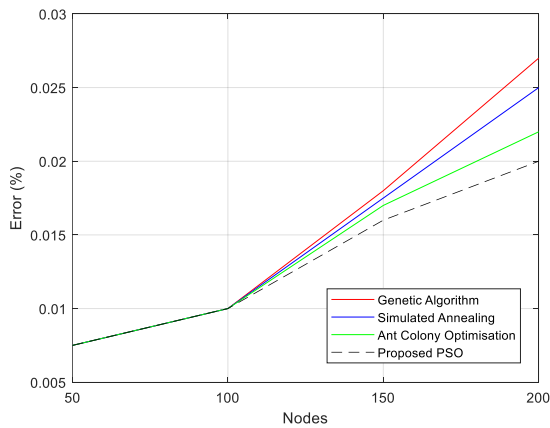


Figure 4: Average Reconstruction Error Error for Sorted Signal

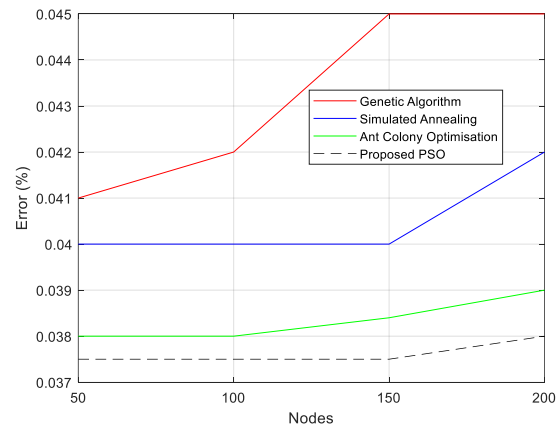


Figure 5: Average Reconstruction Error Error for Unsorted Signal

The reconstruction of sensor values at the base station can be shown in Figure 3, which demonstrates how diverse and noisy data can be employed. This illustrates one of many potential outcomes. Compressive sensing allows for the reconstruction of the initial sensor signal with a significantly reduced degree of error when more measurements are obtained than with traditional sensing methods. The same pattern can be seen in Figure 4, which indicates that the tendency for the reconstruction error to grow can be seen when DCT compression is done in an environment that is noisy. The number of sensor readings that are obtained results in a rise in the error rate, which is proportional to the number of readings shown in figure 5.

This study aims to propose a method for the efficient collecting of data that makes use of adaptive multi-hop routing in conjunction with block-wise compression. This method will be presented as the result of this investigation. The natural signals are related to one another in space, which causes the readings from the sensors to have a sparse distribution as a result.

Compressive sensing may now be utilised in a manner that results in a data collection process that is more efficient. In contrast to the studies that came before it, this one takes use of clustered networks and compressive sensing in its data collection. This indicates that each node in the network is responsible for communicating its observations to the cluster head node. To recover the sensing, cluster heads must first be formed from the measurements received through the use of compressive sensing. Once these cluster heads have been produced, they must then be sent to the base station with multi-hop routing.

It has been demonstrated that the PSO algorithm is capable of significantly lowering the amount of infectious data transfer while simultaneously improving energy efficiency through compressive sensing performance in optimal clustering. This ability of the PSO algorithm to achieve these two goals simultaneously has been referred to as optimal clustering. Because of this method, the network can consume significantly less energy than it would have in any other scenario.

According to the research, one factor that contributes to determining the optimal size for the cluster is having a signal that is relatively weak. When actual sensor readings are used within a cluster, using the discrete cosine transform (DCT) as a sparse basis for clustered compressive sensing enables larger clusters to be deployed with decreased power consumption. This is made possible by employing the DCT as a sparse basis for clustered compressive sensing. The utilisation of DCT as a sparse foundation enables this to become a reality. This approach demonstrates that the one supplied is more trustworthy, effective, and efficient than the customary way things are done, as compared to the typical method.

5. Conclusions

A cluster-based compressive sensing strategy that makes use of an amoeba model has been presented as a potential choice that can bring the clustering procedure to a successful conclusion. After that, the data is sent over a multi-hop network by way of a cluster head that is located at the geographic center of the network. The size of each cluster is used as a basis for estimating the size of the block that corresponds to it in the block diagonal matrix. This is done by comparing the size of each cluster to the size of the block. Power consumption in underwater sensor networks that are structured around compressive sensing while considering the impact that data sparsity has on the efficiency of the underlying sensing algorithm.

Power consumption in underwater sensor networks that are structured around compressive sensing. When the base station is situated in the precise geographic middle of the area that is being sensed, the system functions at the utmost level of effectiveness that it is capable of. The study generates the optimal number of clusters required to accomplish the goal of achieving minimal UASN power consumption if you follow this technique.

Acknowledgement

This research was not funded by any grant.

References

- [1] Wang, S., Lin, Y., Tao, H., Sharma, P. K., & Wang, J., "Underwater acoustic sensor networks node localization based on compressive sensing in water hydrology", *Sensors*, 19(2019): 4552.

- <http://doi.10.3390/s19204552>
- [2] Wang, C., Shen, X., Wang, H., & Mei, H., “Energy-efficient collection scheme based on compressive sensing in underwater wireless sensor networks for environment monitoring over fading channels”, *Digital Signal Processing*, 127(2022): 103530.
<https://doi.org/10.1016/j.dsp.2022.103530>
- [3] Arunkumar, J. R., Anusuya, R., Sundar Rajan, M., & Ramkumar Prabhu, M., “Underwater wireless information transfer with compressive sensing for energy efficiency”, *Wireless Personal Communications*, 113 (2020):715-725.
<https://doi.org/10.1007/s11277-020-07249-7>
- [4] Bhaskarwar, R. V., & Pete, D. J., “Energy efficient clustering with compressive sensing for underwater wireless sensor networks”, *Peer-to-Peer Networking and Applications*, 15(5), (2022): 2289-2306.
<https://doi.org/10.1007/s12083-022-01346-z>
- [5] Chen, X., Xiong, W., & Chu, S., “Two-Tier PSO Based Data Routing Employing Bayesian Compressive Sensing in Underwater Sensor Networks”, *Sensors*, 20(20), (2020): 5961.
<https://doi.org/10.3390/s20205961>
- [6] Chen, X., & Xiong, W., “Compressed Sensing based Multi-Hop Data Routing in 3D Underwater Wireless Sensor Networks”, *15th International Symposium on Medical Information and Communication Technology (ISMICT), IEEE* (2021) :7-12.
<https://doi.org/10.1109/ISMICT51748.2021.9434915>
- [7] Erdem, H. E., Yildiz, H. U., & Gungor, V. C., “On the lifetime of compressive sensing based energy harvesting in underwater sensor networks”, *IEEE Sensors Journal*, 19(12), (2019): 4680-4687.
<https://doi.org/10.1109/JSEN.2019.2900427>
- [8] Lin, C., Han, G., Wang, T., Bi, Y., Du, J., & Zhang, B., “Fast node clustering based on an improved birch algorithm for data collection towards software-defined underwater acoustic sensor networks”, *IEEE Sensors Journal*, 21(22), (2021):25480-25488.
<http://doi.org/10.1109/JSEN.2021.3055948>
- [9] Lin, Y., Tao, H., Tu, Y., & Liu, T., “A node self-localization algorithm with a mobile anchor node in underwater acoustic sensor networks”, *IEEE Access*, 7, (2019): 43773-43780.
<http://doi.org/10.1109/ACCESS.2019.2904725>
- [10] Xian, J., Wu, H., Mei, X., Zhang, Y., Chen, X., Zhang, Q., & Liang, L., “Novel Energy-Efficient Opportunistic Routing Protocol for Marine Wireless Sensor Networks Based on Compressed Sensing and Power Control”, *Journal of Ocean University of China*, 21(6), (2022):1504-1516.
<https://doi.org/10.1007/s11802-022-5128-6>
- [11] Thangarasu. G, Dominic. P.D.D, Subramanian. K, “An Efficient Energy Consumption Technique for Integrated Internet of Things Operations”, *17th IEEE Students Conference on Research and Development, University Technology PETRONAS, Malaysia*, 15-17 October (2019): 45-48.
<http://doi.org/10.1109/SCORED.2019.8896238>
- [12] Singh. A, “Managing the Salinization and Drainage problems of Irrigated Areas Through Remote Sensing and GIS Techniques”, *Ecological Indicators* 89(2018):584-589.
<https://doi.org/10.1016/j.ecolind.2018.02.041>

- [13] Thangarasu. G, Dominic. P.D.D, Subramanian, “Efficient Energy Usage Model for WSN-IoT Environment”, *International Conference on Computational Intelligence, University Technology PETRONAS, Malaysia*, (2020): 252-255.
<http://doi.org/10.1109/ICCI51257.2020.9247795>
- [14] Deqing. W, Ru. X, Xiaoyi. H and Wei. S., “Energy-Efficient Distributed Compressed Sensing Data Aggregation for Cluster-Based Acoustic Sensor Networks”, *International Journal of Distributed Sensor Networks*, 21(22), (2016): 01-14.
<http://doi.org/10.1155/2016/8197606>
- [15] Lin. C, Han. G, Qi. X, “Energy Optimal Data Collection for Unmanned Aerial Vehicle-Aided Industrial Wireless Sensor Network-Based Agriculture Monitoring System: A Clustering Compressed Sampling Approach”. *IEEE Transaction on Industrial Informatics*, 17(06), (2022): 4411-4420.
<http://doi.org/10.1109/TII.2020.3027840>

IMPROVED ERROR FREE CLUSTERING IN COGNITIVE RADIO NETWORKS USING MODIFIED PARTICLE SWARM OPTIMISATION

Gunasekar Thangarasu ^{1*}
Kesava Rao Alla ²

¹Department of Professional, Industry Driven Education, MAHSA University, Saujana Putra, Malaysia

²Chancellery, MAHSA University, Saujana Putra, Malaysia

Abstract: *The goal of this paper is to develop and simulate a coverage and localization model that minimizes power consumption while maintaining maximum coverage. To achieve this, the sensor nodes are divided into multiple groups, with each group responsible for a specific aspect of monitoring. This approach is effective because there are more sensors available than strictly required, enabling them to be utilized efficiently. To optimize energy savings, the ad hoc node is strategically positioned to take advantage of the most effective energy-saving opportunities. By doing so, the capacity of cognitive radio network (CRN) sensor nodes to conserve energy can be improved significantly. Overall, this model offers an efficient and practical approach to power conservation while maintaining optimal coverage in CRN sensor networks.*

Keywords: *Energy consumption, Cognitive radio, Routing, PSO*

1. Introduction

Unlicensed users, also known as Secondary Users, can discover empty channels and take advantage of them opportunistically by employing Cognitive Radio (CR) [1]. This is possible provided that the interference to Primary Users (PUs) is low enough. This is only doable if the amount of interference from the PU is sufficiently minimal. User units, also known as SUs, in a distributed cognitive radio network can connect with one another despite the lack of a central hub or base station in the network.

It is important to keep in mind that the formation of clusters out of the nodes that make up a network could potentially improve the overall performance of the system. Clustering is utilised in CR networks because to the many benefits that it gives, all of which are vital to the successful performance of CR operations. These benefits include help for cooperative tasks such as channel sensing and access, scalability, and stability [2].

Further benefits include the ability to scale. It should come as no surprise that clustering is a significant component of CR networks that calls for additional investigation. This is since the channel availability of each SU, which is also referred to as white spaces, varies depending on the level of activity that is taking place at the PU, the time of day, and the location of the PU. In other words, the level of activity at the PU determines how many white spaces are available. The tried-and-true ways of clustering, such as utilising the lowest ID [3] and the greatest node degree [4], have been the focus of several novel clustering algorithms that have been suggested as potential alternatives. By using the method of clustering with minimal identifiers, the node that is chosen to be the leader of the cluster is the one that has the fewest identifiers overall. When the maximum node degree clustering algorithm is used, on the other hand, the node that

is chosen to be the leader of the cluster is the one that is connected to the greatest number of other nodes. Traditional clustering algorithms, such as those used in wireless ad hoc and sensor networks [5–7], are not appropriate for use with CR networks for several reasons, the most important of which are as follows:

To begin, the algorithms that regulate CR networks are not designed to develop in such a way that they can account for the constantly shifting channel conditions. Second, there is the chance that the algorithms won't be able to achieve the goals that are unique to CR systems. This is a risk that must be taken into consideration. For instance, if a cluster does not have enough white spaces, certain nodes in the cluster would not be able to communicate with the cluster head.

It is possible that the strategies will not increase application-specific network performance, such as achieving a lower error probability in channel sensing outcomes, or CR-specific network performance, such as increasing the number of common channels in a cluster. Both outcomes would fall under the category of the possibility of not increasing network performance.

2. Cognitive Radio Network

Conventionally, the radio spectrum is segmented into narrower ranges of frequency bands, each of which might be licenced or unlicensed for individual use. This is the standard protocol for the distribution of spectrum (also called channels). The government uses auctions to distribute licenced channels so that public utilities can have access to channels that are available to them exclusively. It is prohibited for SUs to use any of the approved channels, including the widely deployed IEEE 802.11 wireless communication protocols. Instead, to watch television without paying for it, they tune in to stations that are not authorised. With cognitive radio, SUs are able to monitor the available radio frequency spectrum and make use of white spaces, all while causing just a minimal level of interruption to PUs. This is made possible via cognitive radio. The total amount of available bandwidth across all SUs needs to be increased, and therefore, the utilisation of radio spectrum, which is a limited resource for wireless networks, needs to be maximised. This can only be accomplished by increasing the total amount of available bandwidth [8-11].

Every sensor node that is included in a CRN is subject to the same significant constraints that are imposed on WSNs. These restrictions apply to both the hardware and the energy that it has. A good number of CRNs have made it one of their highest priorities to cut down on the total amount of energy that they consume. One of the most notable distinctions between CRNs and conventional CR networks is that, in the former, data packets can only be generated by nodes in the network that have detected an event and then sent to a sink node, whereas in the latter, data packets can be generated by any node in the network and then sent to a base station. This is one of the most important differences between the two types of networks. This is among the most significant distinctions that can be made between the two categories of networks [12].

Clustering is a means of regulating network topology that is a strategy for boosting the performance of a network. This is accomplished by merging individual nodes into logical units that are referred to as clusters. Clustering is a method of controlling network topology. The cluster structure that is produced is in turn dependent on aspects such as the position and channel availability of the nodes in the underlying network, often known as white spaces [9].

A cluster is made up of several different types of nodes, the most important of which are the central hub (CH), member nodes, and gateway nodes. Each of these nodes performs a specific role that is essential to the cluster overall architecture. In a CR, a CH job is to act as a local

processing centre for tasks like channel sensing and routing, both of which are very necessary to the functioning of CRs. Each linked node has an associated connected node or connected nodes (CNs), which can range from one to multiple (CH). To demonstrate how this works, CH has connections to all the member nodes [13].

As part of the cluster intra-cluster communications, the CH and the member nodes maintain regular communication with one another. The gateway nodes of a cluster are the nodes that are positioned in the peripheral of the cluster. These nodes are sometimes referred to as edge nodes. They are responsible for facilitating communication across clusters. Gateway nodes can take in messages sent from other clusters because of their privileged position. One example of this would be a gateway node that is connected to a CH, which streamlines the process by which clusters can communicate with one another and makes it easier for clusters to work together.

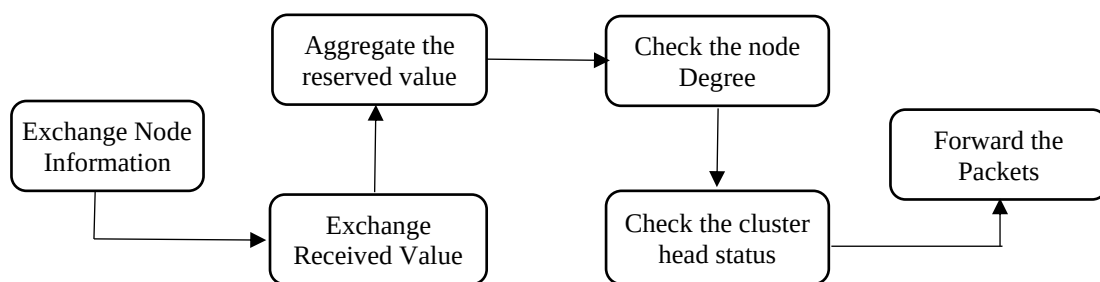


Figure 1: Proposed Methodology for CRN

It is common practise for neighbouring clusters to make use of different channels in CR networks, a gateway node may frequently need to change the operational channel that it is using. This is because it is common practise for neighbouring clusters to make use of different channels. The CHs and gateway nodes make up the backbone of the network in this setup. The backbone of the network is the component that oversees connecting the BS to the other nodes in the network.

The process of establishing a cluster, which frequently involves selecting a CH and having member nodes join together, is the stage where the various clustering algorithms diverge the most from one another. If a node is unable to immediately locate an already existing cluster, it will start the process of forming a new cluster as soon as possible. In order to fulfil global goals such as increasing the overall network throughput and the number of channels that are shared, nodes in the surrounding area will broadcast information about themselves, such as their channel availability lists.

Important clustering criteria, such as the availability of channels, are taken into mind before a decision on a CH is made. Both examples show that a CH has the highest node degree level. Both illustrations illustrate the fact that a CH possesses the maximum possible node degree level. CH possesses a greater node degree level, a reduced number of hops, and fewer channel switches on the path that leads from member nodes to the CH. Moreover, the total number of hops taken is cut down.

The process of a node deciding which cluster it will join is one of the steps involved in the process of the node becoming a member of the network. The number of common channels that

a node has with the cluster will be one of the variables that will be taken into consideration in making this decision [11]. If there are more common channels among the players who make up a cluster, the re-clustering process will not take place when PU operations restart. It is important to consider the likelihood that a node, when entering a cluster, may limit the number of shared channels, which could result in the system being unstable.

3. Proposed Method

A level-based clustering technique is excellent for setting the threshold that is a factor in the process of cluster creation since it delivers more exact results. Using a method that is based on levels allows for this, which is one of its many advantages. It is important for a node to generate a random number for that node to take part in the competition for CH. This competition is for the control node. If the evaluated threshold, $T(i)$, is greater than the random number, the random number is first converted into CH using the following method [14-15]:

$$T(i) = \frac{P_{opt}}{1 - p_{opt} \left(r \cdot \text{mod} \left(\frac{1}{P_{opt}} \right) \right)} * \frac{E_i(r)}{E_{avg}(r)}$$

r - present iteration in CRN,

$E_i(r)$ - energy of a node i .

E_{avg} - average residual energy

$$E_{avg} = \frac{\sum E_i(r)}{N} \text{ and}$$

N - total nodes.

The energy consumption is defined:

$$E_n T_{rx}(L, D) = \begin{cases} LE_{n(elec)} + L\epsilon_{fres} D^2, & D < D_0 \\ E_{n(elec)} + L\epsilon_{m_{pat}} D^4, & D \geq D_0 \end{cases}$$

where,

L - packet size.

ϵ_{fres} - energy loss.

$\epsilon_{m_{pat}}$ - multipath energy loss.

D_0 - threshold distance

$$D_0 = \sqrt{\frac{\epsilon_{fres}}{\epsilon_{m_{pat}}}}$$

3.1 Neighbor Hood Estimation

The collection of all the various possible solutions for which at least one of the transformations is valid is referred to as the neighbourhood of S . This is the meaning behind the term neighbourhood of S . If a solution can satisfy each and every one of these prerequisites, then we can confidently state that it is located pretty close to S . It is possible to build a wide range of communities, each of which has the potential to be constructed with a certain mode of

transportation in mind. This can be done because it is feasible to build a large variety of communities.

When looking at production, the most productive sector is the one that produces the most with the fewest obstacles to overcome, and this is especially true when looking at production on an individual level. In addition, we examine the effectiveness of the extra-potential neighbourhoods and rate them based on their potential from most successful to least successful.

3.2 Node Deployment

If a solution can be found that satisfies the needs of these networks for connectivity and coverage, then the issue of deploying nodes in mesh networks can be overcome, and the mission can be finished successfully. If you want to provide connectivity and coverage with a CRN, the most effective strategy is to utilise a plan that makes use of a manner that asks for as few sensors as is practically possible, as this is the most cost-effective approach. Although it is not an ideal solution, the ant colony method offers a workable answer to the networking and coverage difficulties brought on by the deployment of the nodes.

On the other hand, this method does not guarantee success in every situation. Nonetheless, there are a great deal of limitations that must be taken into account, such as more stringent standards and more complicated ecosystem designs. The negative aspects that will be covered in the following section provide a striking contrast to the positive aspects that have already been covered. There have been several attempts made to increase the reliability of the ant colony strategy by utilising a wide variety of management and optimization strategies. Some of these attempts have been successful, while others have been unsuccessful.

It is not yet dependable because it exhibited early or late convergence, in addition to several other irregularities that were observed during testing. This is the case although the strategy based on ant colonies has been subjected to a variety of control and optimization procedures. Because it uses a culture algorithm as a guide for the ant-colony algorithm, this method is able to improve the robustness of the ant-colony algorithm and speed up the search process. Ant-colony algorithms are commonly used in search engine technology. The insertion of the improved ant-colony algorithm into the culture algorithm population region through the use of the two-layer culture algorithm evolution technique is the conceptual backbone of the ACO algorithm.

The algorithm is now in a stronger position to deal with any problems that may appear in the future. In this subject of computer science study, an algorithm is used to find the best possible answer, in addition to any other possibilities that, given the characteristics of the problem, may be made to perform successfully. Throughout the course of time, there will be an increase in both the total number of individuals living in space as well as the variety of people who make up that population.

3.3 Routing Model

The value can be continually changed by making use of fuzzy sets and membership criteria, which is a beneficial quality to have. This can be accomplished through the usage of fuzzy sets. When performing the computation for the change in inertia, it is possible to utilise either the current value or the fitness of the best particle as the starting point. These two choices are equivalent in their applicability. The characteristics of the search space have a key role in

establishing the level of complexity associated with the optimization problem and, as a direct result of this, the value that is assigned to the parameter.

3.4 Construction Factor

The concept of the constriction coefficient was initially established because of an eigenvalue analysis of computational swarm dynamics during the early phases of its development. This study was carried out to better understand the behaviour of the system. By applying the updated formula, we can determine the current velocities, and they are as follows:

$$\frac{2v}{\left| 2 - \Omega - \sqrt{\Omega(\Omega - 4)} \right|}$$

$$v_x[i] = \chi \times (v_x[i] + \Omega_1 \times \text{rand}() \times (pBest[i] - presentx[i]) + \Omega_2 \times \text{rand}() \times (pBest[i][gbest] - presentx[i]))$$

$$\Omega = \Omega_1 + \Omega_2$$

Ω_1 – product consisting of c1 and r1

Ω_2 - product consisting of c2 and r2

The average particle speed decreases as the number of repeats grows, swarm convergence may be guaranteed under the operational conditions $v \in [0, 1]$, and 4. This is because the number of repetitions increases. You could change how far the swarm will search for its prey. The inspection of the particles at each stage of their passage through the search area should be the top priority for this endeavour.

Since the value is so near to 1, this becomes a very critical consideration. When dealing with a problem that has a higher degree of complexity, the general rule of thumb is that slowing down the convergence rate will result in a solution that has a higher quality, and this can be achieved by reducing the convergence rate. Yet, when the problem at hand is uncomplicated, convergence occurs more rapidly, even though there is still room for a certain amount of variation in the quality of the answer (v).

This method of limiting velocities has the same impact on evaluating the quality of the solution throughout all the neighbourhoods of the search space as the variance in the inertia weight does. This is because both have the same effect on how well the solution works. to arrive at our destination in a more expedient manner, clamping and velocity restriction might be employed in conjunction with one another. The intended impact might be achieved by combining both components into one effort.

3.5 Cognition Velocity

This investigation is one of the earlier ones that investigated the effect that attractor topology has on the precision of swarm trajectory map updates. The cognitive model, in contrast to the standard PSO, which also considers monetary and physical concerns, is exclusively concerned with mental considerations, as stated in Equation 1. This distinction may be seen in the fact that the cognitive model focus is limited to mental factors (1).

$$v_{t+1}[i] = (v_t[i] + C1 \times \text{rand}() \times (pBest[i] - presentx[i]))$$

The cognitive model will do a local search in the area where the swarm members are present if both the acceleration component and the upper velocity restrictions are low. As a direct result

of this, there is an increased likelihood that reports of results that are only satisfactory will be produced. Its limited exploratory potential is mostly to blame for its sluggish rate of convergence, which can be characterised in the following manner: The social model, on the other hand, positions persons as the major component of the solution by putting an emphasis on them as the primary focus of attention.

$$vt+1[i][j]=(vt[i][j]+C1\times rand() \times (pBest[i][j]-presentx[i][j]))$$

They are pulled to the local best in the feasible zone, the particles in the exploratory model converge more quickly with a bias towards discovery when compared to the traditional model. This is because the particles in the exploratory model are drawn to the feasible region.

4. Results and Discussions

This section provides an approachable high-level summary of the benchmarking tests that were carried out using PSO and some of its modifications. The tests were carried out to compare PSO to other systems. The tests were carried out with the purpose of contrasting PSO with other systems that are very comparable to it. It is standard practise to carry out statistical significance tests at a high enough level of confidence before drawing any conclusions regarding the performance of the algorithms.

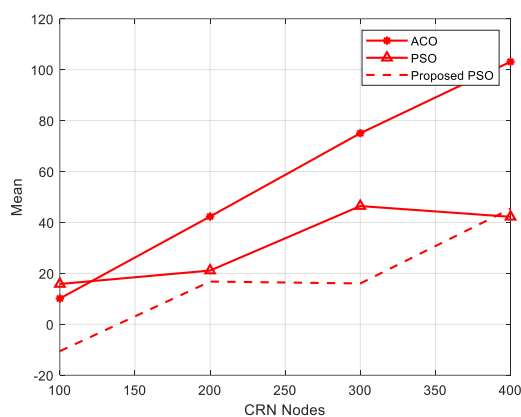


Figure 2: Mean of Fixed CRN

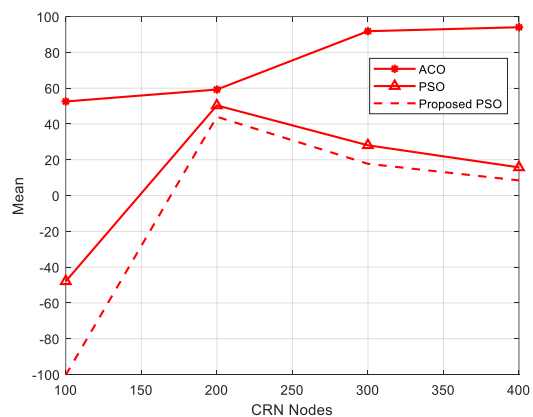


Figure 3: Mean of Mobile CRN

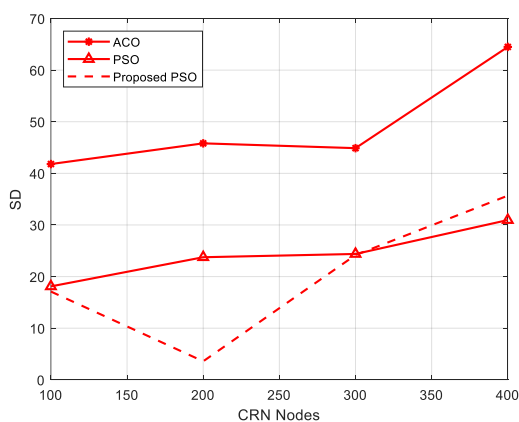


Figure 4: SC of Fixed CRN

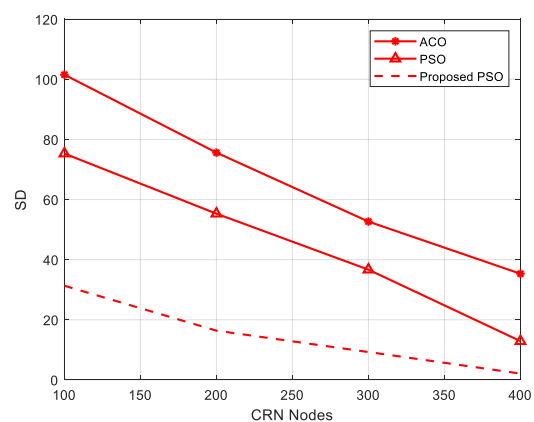


Figure 5: SC of Mobile CRN

To provide a bird-eye view and a perspective grounded in common sense on the PSO benchmarking study, the findings from figures 2 - 5. It is of statistically significant with a

degree of confidence that is suitably high before arriving at any inferences regarding the performance of the algorithms. These tests should be carried out with the utmost certainty.

5. Conclusions

The quality of the solutions that are generated by metaheuristics such as PSO is significantly influenced by the parameters that are utilised during the training process of the algorithm. If the objective function is sufficiently complex, the search space will quickly become dominated by local optimal solutions if the fundamental PSO is not extensively adjusted to account for the modalities of the goal function. This will occur if the fundamental PSO is not adjusted to account for the modalities of the goal function. This is going to take place no matter how drastically the PSO is changed. PSO methods can be used to produce good solutions for multimodal functions in both static and dynamic circumstances. Even though the quality of these solutions rapidly diminishes as the dimensionality of the problem increases, PSO methods can be used to accomplish this goal.

Acknowledgement

This research was not funded by any grant.

References

- [1] Bhatti, D. M. S., Ahmed, S., Chan, A. S., & Saleem, K., "Clustering Formation in Cognitive Radio Networks using Machine Learning", *AEU-International Journal of Electronics and Communications*, 114 (2020): 152994.
<https://doi.org/10.1016/j.aeue.2019.152994>
- [2] Hossain, M. A., Schukat, M., & Barrett, E., "Enhancing the Spectrum Sensing Performance of Cluster-based Cooperative Cognitive Radio Networks via Sequential Multiple Reporting Channels", *Wireless Personal Communications*, 116 (2021):2411-2433.
<https://doi.org/10.1007/s11277-020-07802-4>
- [3] Zhao, F., Li, S., & Feng, J., "Securing Cooperative Spectrum Sensing Against DC-SSDF Attack using Trust Fluctuation Clustering Analysis in Cognitive Radio Networks", *Wireless Communications and Mobile Computing*, 2019.
<https://doi.org/10.1155/2019/3174304>
- [4] Prajapat, R., Yadav, R. N., & Misra, R., "Energy-Efficient K-hop Clustering in Cognitive Radio Sensor Network for Internet of Things", *IEEE Internet of Things Journal*, 8(17), (2021): 13593-13607.
<https://doi.org/10.1109/JIOT.2021.3065691>
- [5] Sharma, G., & Sharma, R., "Energy Efficient Collaborative Spectrum Sensing with Clustering of Secondary Users in Cognitive Radio Networks", *IET Communications*, 13(8), (2019): 1101-1109.
<https://doi.org/10.1049/iet-com.2018.5109>
- [6] Zheng, M., Chen, S., Liang, W., & Song, M., "NSAC: A Novel Clustering Protocol in Cognitive Radio Sensor Networks for Internet of Things", *IEEE Internet of Things Journal*, 6(3), (2019): 5864-5865.
<https://doi.org/10.1109/JIOT.2019.2898166>
- [7] Kumar, S., & Singh, A. K., "A Localized Algorithm for Clustering in Cognitive Radio Networks", *Journal of King Saud University-Computer and Information Sciences*, 33(5), (2021): 600-607.
<https://doi.org/10.1016/j.jksuci.2018.04.004>

- [8] Vimal, S., Khari, M., Crespo, R. G., Kalaivani, L., Dey, N., & Kaliappan, M., “Energy Enhancement using Multi-Objective Ant Colony Optimization with Double Q Learning Algorithm for IoT Based Cognitive Radio Networks”, *Computer Communications*, 154, (2020): 481-490.
<http://doi.org/10.1109/JSEN.2021.3055948>
- [9] Stephan, T., Al-Turjman, F., Joseph, K. S., Balusamy, B., & Srivastava, S., “Artificial Intelligence Inspired Energy and Spectrum Aware Cluster Based Routing Protocol for Cognitive Radio Sensor Networks”, *Journal of Parallel and Distributed Computing*, 142, (2020): 90-105.
<https://doi.org/10.1016/j.jpdc.2020.04.007>
- [10] Ghaznavi, M., & Jamshidi, A., “A Low Complexity Cluster Based Data Fusion to Defence Against SSDF Attack in Cognitive Radio Networks”, *Computer Communications*, 138, (2019):106-114.
<https://doi.org/10.1016/j.comcom.2019.03.007>
- [11] Thangarasu. G, Dominic. P.D.D, Subramanian. K, “An Efficient Energy Consumption Technique for Integrated Internet of Things Operations”, *17th IEEE Students Conference on Research and Development, University Technology PETRONAS, Malaysia*, 15-17 October (2019): 45-48.
<http://doi.org/10.1109/SCORED.2019.8896238>
- [12] Pari. D and Natarjan. J, “Defence Against SSDF Attack and PUE Attack in CR-Internet of Vehicles for Millimetre Wave Massive MIMO Beamforming Systems”, *Symmetry*, 14,12(2022):2472.
<https://doi.org/10.3390/sym14122472>
- [13] Thangarasu. G, Dominic. P.D.D, Subramanian, “Efficient Energy Usage Model for WSN-IoT Environment”, *International Conference on Computational Intelligence, University Technology PETRONAS, Malaysia*, (2020): 252-255.
<http://doi.org/10.1109/ICCI51257.2020.9247795>
- [14] Feng. J, Zhang. M, Xiao. Y and Yue. H., “Securing Cooperative Spectrum Sensing Against Collusive SSDF Attack Using XOR Distance Analysis in Cognitive Radio Networks”, *Sensors*, 2018, 18(2): 370
<http://doi.org/10.3390/s18020370>
- [15] Venkatesan. P and Shanmughavel. S, “Attack Detection and Securer Data Transmission in Cognitive Radio Networks Using BMHHO-ENN and SHA2-RSA”. *Soft Computing*, 2021, 1(5): 175-187.
<https://doi.org/10.1007/s00500-021-06462-1>

BLOCKCHAIN-BASED DEEP LEARNING FOR SUSTAINABLE AGRICULTURAL SUPPLY CHAIN MANAGEMENT

Kesava Rao Alla^{1*}
Gunasekar Thangarasu²

^{1*}Chancellery, MAHSA University, Saujana Putra, Malaysia

²Department of Professional, Industry Driven Education, MAHSA University, Saujana Putra, Malaysia

Abstract: Food supply chain (FSC) is an important part of the food supply chain. It is essential to establish food supply chains that are open to the public, accountable for their actions, and available in real time. The development of blockchain technology has led to an increase in the amount of data that is passed between customers and businesses, as well as the data that are passed between them. Blockchain technology is a new type of information technology that has the ability to be decentralized, safe, and trusted, which makes it an excellent option for storing sensitive data. The purpose of this study is to evaluate how blockchain with deep learning is used to find the quality evaluation of food supply chain technology. The use of BC has improved the accuracy of food traceability, while the utilization of Deep Random Forest (DRF) has boosted the efficacy of computing and shortened the reaction time. The research compare the quality evaluation system that is based on BC-DRF with some of the most common existing methods in terms of accuracy, reaction time, and sensitivity, when applied to a variety of block sizes.

Keywords: Blockchain, Deep learning, Sustainable Agriculture, Supply Chain Management

1. Introduction

Instead of depending on a centralized server or any other authoritative entity, the management of the blockchain data is handled by a decentralized network of computers. Over the course of the past few months, the blockchain technology has been integrated into a variety of new applications [1]. If two parties to an agreement cannot be trusted, for instance, the conditions of the agreement can still be carried out and enforced with the use of a smart contract that is based on blockchain technology. It may reduce the amount of time and effort needed to negotiate the terms of a contract [2].

Another area where blockchain technology could potentially be applied is in the management of supply chains. Walmart has developed a blockchain system by employing Hyperledger Fabric. This is done so that the company can evaluate the standard of the food supply that it receives. There is a significant amount of interest being shown in blockchain technology coming from other sectors [3].

In a world that is becoming increasingly dominated by digital technology, the manual processing of vast amounts of data is inefficient due to the length of time it takes. The application of machine learning as a strategy for the processing of a diverse range of data sources is gaining widespread traction. This is mostly attributable to the fact that it is able to

automatically learn from its past mistakes and better itself as a result of doing so. Training supervised learning algorithms with data that has been tagged in some way is possible [4]. The fact that blockchain technology is decentralized, safe, and trusted makes it an excellent option for storing sensitive data. Large volumes of data can be processed and analyzed by machine learning systems. Because of this, a substantial number of recent studies have studied the possibility of combining blockchain technology with machine learning in order to create real-time systems that are secure, efficient, and friendly to the environment. For instance, the procedures involved in accounting by employing the technologies of machine learning and blockchain. In addition, the blockchain technology that has been improved by machine learning is utilized in the medical industry [5].

The control over a series of flows, beginning with production and ending with consumption, is required for the management of the supply chain. Getting your hands on the fresh components is the initial stage in the process, and handing off the finished good to the customer is the very last thing to do.

In order for these supply chains to become more effective as a whole, new mechanisms for increased cooperation, resilience, and responsiveness will need to be built. The proliferation of blockchain technology has directly contributed to a increase in the amount of supply chain management in agricultural industry.

2. Related Works

The advancement of digital and information technology is causing a commotion in the interactions that take place between customers and businesses, as well as the data that is passed between them. These advancements are due, in large part, to technological innovations such as smartphones, wearables, drones, computers connected to the IoT, machine learning and artificial intelligence. These cutting-edge technologies and algorithms can be helpful to the food supply chain in a variety of different ways, and the ones that have been mentioned here are only few of those ways. One such example is keeping a close eye on the weather and conducting research into the ways in which it can affect the economy of the agricultural sector [6]. It can be used in the construction of systems that track the production and distribution of food at every stage, from the farm to the consumer plate, and it can be put to use in both of those processes.

In the context of the food supply chain, the distributed ledger technology known as blockchain is regarded to be a tool that enables transactions to be honest and reliable. The costs of transactions, the governance of supply chains, and the functionality of resources are all considered to have the potential to drop as a result of blockchain technology. These benefits can be seen in the food supply chain. By utilizing blockchain technology as a data format, it is possible to realize data parity in food supply chains that are centered on blockchain technology. Smart contracts can be used to automate procedures such as the payment of vendors, and the many parties engaged in the food supply chain are able to store and share data securely. Additionally, they are able to accomplish a great deal more. In the areas of law and finance, a smart contract can also perform the functions that are traditionally associated with other types of contracts [7].

A strategic tool that has the ability to affect the reorganization of food supply chains as well as the capacities of food organizations is the distributed ledger technology, which is also known

as blockchain. Partners in the supply chain were more dependent not just on one another but also on interactions that were vertically coordinated as a result of the expenses involved with installing blockchain technology at various points along the supply chain. This was a direct result of the costs associated with implementing blockchain technology at various points along the supply chain. It will also help develop alliances and collaboration inside the supply chain, which is a great benefit, so that something to look forward to. In addition to the aforementioned advantages, the use of the blockchain technology will also foster the development of skills that are beneficial to the administration of chain-wide innovation [8].

With the assistance of technologies such as blockchain, sensors, the Internet of Things (IoT), the cloud, and artificial intelligence (AI), it is possible to establish a food supply chain that is open to the public, accountable for its actions, and available in real time. Both blockchain technology and the Internet of Things (IoT) have the potential to have an impact on a wide range of domains, including but not limited to the management of food supply chains, automated payments, quality control, and provenance [26]. Blockchain technology is a distributed ledger that can be used to record transactions in a decentralized manner [9].

Processing and analysis of massive datasets made available by blockchain and AI applications are necessary for the production, distribution, and consumption of food in a sustainable manner. This is a necessary step in the process of making the entire food supply chain more sustainable. The implementation of blockchain technology within the FSC enables a transparent and trustworthy traceability system, which is vital for the upkeep of public health in the event that food, fruit, or vegetables are suspected to be traced back to their place of origin [10]-[11]..

3. Proposed Method

The purpose of this study is to evaluate how many different components of blockchain technology and its potential applications in food supply chain that contributes to the accomplishment of the Sustainable Development Goals (SDGs) as illustrated in Figure.1.

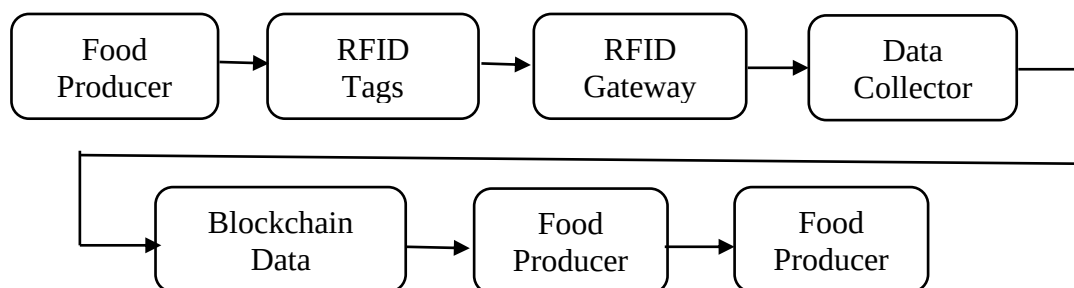


Figure 1: Proposed BC-SCM

3.1 Data Acquisition

Let pretend you have a dataset A with n food samples, and that each sample has information about the food color, turbidity, fat, odor, taste, temperature, and pH. Let also pretend you have a dataset B with n food samples, and that dataset B also has n food samples. After being subjected to normalization, the dataset can be interpreted in the form

$$A = \{A_1, A_2, \dots, A_n\}$$

where, A_l - l^{th} sample after the normalization process.

3.2 Data Pre-processing

The dataset A_l goes through the pre-processing step, which is when the Z-score normalization that was utilized in order to clean the data set is applied. After that, the data set A is ready for analysis. In artificial neural networks, the Z-score is typically utilized as the first stage of the process. It has a significant computational capability and takes into account the mean as well as the variability of the raw data. Additionally, it displays the results.

The research normalizes the data by taking into account the mean as well as the standard deviation of the variables that were input, with the goal of having the mean and the standard deviation of all of the characteristics equal to zero and one, respectively. We do this so that the mean and the standard deviation of all of the characteristics are the same. The Z-score method of normalizing can be defined as follows:

$$b_l = \frac{A_l - \mu_l}{\sigma_l}$$

Here,

μ_l - arithmetic mean and

σ_l - variance of A_l and

b_l - normalized data.

3.3 Feature Selection

In order to obtain the most value possible from the pre-processed data, the normalized data b_l are put through a feature selection technique. This is done in order to maximize the potential for the data. In this particular scenario, features are retrieved with the assistance of the Canberra distance, which measures the amount of distance that separates a feature and a class label. The distance to Canberra is not overly susceptible to being changed by outlying data points. The following equation can be utilized in order to calculate the distance between the starting point and Canberra.

$$CD(g, h) = \sum_{j=1}^m \left| \frac{g_j - h_j}{g_j + h_j} \right|$$

where,

g - candidate feature and

h - target, and

m – vector dimension.

The chosen features f_s are recorded into the DRN so that they can be considered for determining the overall quality of the food.

$$f_s = \{f_1, f_2, \dots, f_k\}$$

3.4 Decision Tree

The decision tree approach (DT) is a type of predictive learning that combines the opinions of subject matter experts regarding the intended outcome of an investigation with actual data pertaining to the phenomena that are the focus of the research. The goal of the inquiry is to determine whether or not a certain phenomenon exists. Learning is the method that is implemented in this technique in order to locate a DT within the dataset.

In the subject of spatial data mining, this methodology is applied rather frequently in research projects. The probable range of values for each target variable is calculated based on the values of the input variables, and the leaves of the decision tree each reflect a different range of

possible outcomes for each target variable. One can unearth a DT that had been concealed by concealing it before by separating a source set according on the results of an attribute value test. It is of the utmost importance that this method be followed consistently when deriving each and every subset possible.

When it is possible to assign a single classification to each sample in the subset that was produced by the approach, the recursive operation comes to an end. This can happen either when it is no longer necessary to perform additional splitting or when it is possible to do so. These models are useful for both predictive modeling and classification, and this is due to the fact that the rules that are produced by DTs are simple enough for humans to understand.

The finished DT structure need to have an extremely restricted amount of movable parts. In order for DTs to accurately forecast new data, they need to have the ability to unearth the information that they have gained from the dataset in a manner that is consistent throughout the process. The great majority of DTs are amenable to modification or improvement when new training data are incorporated into the learning process.

There are many subgroups that fall under the umbrella of DTs. A few different DTs each produce output in the form of a separate list of numbers. CARTs allow for the development of results in both the classification and regression domains, and are therefore commonly referred to as classification categorization tree. The CART method was used throughout the duration to analyze the data and produce suggestions for selecting a path of action. The initial node, sometimes known as the root node, was the location where all of the data for the CART model was saved.

The final step involved applying the splitter variable that possesses the highest degree of purity and homogeneity across all of the branches in order to create a split in the root. This was the phase that was responsible for creating the split. At each of the tree branches, a categorization of the variables was carried out until the appropriate node included data that was enough comparable to fit into a single classification. This process would continue until the correct node had all of the data in its storage space.

The leaf nodes provide a representation of the several classes that have developed from the tree. The internal nodes of a network were situated in a location that was midway between the extremes of the network and the network central node. When trying to determine what the splitter variable is, one has access to a number of indices that they can use in their investigation. In order to carry out the analysis on the data, the Gini index, also known as the GI, was utilized. On the other side, it was discovered that the value of the GI was at its highest when each node held data from each and every possible category. When all of the information that was connected to a specific node belonged to the same group, that node was constructed with the highest possible degree of accuracy, and the GI value that was associated with it was 0.

When all of a node variable had been determined, the GI of each one was computed, and the dataset variable that had the shortest GI was selected to serve as the splitter. This process was repeated until all of the variables of a node had been determined. The following formula was applied so that the required GI value could be determined for CART:

$$GINI(p) = \sum_{i=1}^n p_i (1 - p_i)$$

where p_i - likelihood of object placed in a category

3.5 Training of RF using SOM

The two layers that are included in a SOM are referred to as the input layer and the output layer. There is a sizeable portion of neurons found in the output layer that have established complete connections with neurons found in the input layer.

A learning algorithm that concurrently employs unsupervised and competitive learning methodologies is used to teach the SOM. This type of learning is referred to as competitive learning.

The steps that are involved in SOM training are broken down into their component parts in the following condensed summary. The pattern that is used as input to the SOM can be described in mathematical terms as follows:

$$x = [x_1, x_2, x_M]^T$$

where

x - input pattern and

M - size of x .

As a result, the SWVs of the neurons in the output layer are represented by:

$$w_j = [w_{j1}, w_{j2}, w_{jM}]^T$$

where

$w_j, j=1, 2, \dots, p$, - j^{th} neuron SWVs and

p - Kohonen layer's.

In the initial stage of the training process, the SWVs are provided with a set of values that are relatively arbitrary to serve as a starting point. The initial values for the learning phase of the self-organizing map will be quite arbitrary, and the map will utilize these values as its point of departure. Now we will move on to the stage of the self-organizing map known as learning. Within the context of self-organizing maps, this style of education is referred to as competitive learning. This is because the neurons in the system compete with one another to see which neuron will be triggered and go on to become the winning neuron. The neuron that is stimulated will then go on to become the victorious neuron.

The SWV of each neuron is compared to the input pattern, and the winner is the neuron that exhibits the greatest degree of congruence between the two. After that, the SWV parallels are applied to arrive at a conclusion regarding which neuron proved winning. Because of this, we can determine which neuron emerged victorious by applying the following formula:

$$i(x) = \arg \min \|x - w_j\|, j=1, 2, \dots, p$$

where

$\|\cdot\|$ - Euclidean distance and

$i(x)$ - neuron near to x .

$$\|x - w_j\| = \sqrt{\sum_{i=1}^M (x_i - w_{ji})^2}, j=1, 2, \dots, p$$

The Euclidean distance is often used as a measuring stick by those who are comparing and contrasting two different objects. When the value of $\|x - w_j\|$ is reduced, the degree to which the

input x and the SWV w_j correspond to one another increases. Following this step, the distances between the active node x in the network and each of the network output neurons are compared in order to determine which neuron in the network.

The neurons that are successful are those that are situated in the part of the brain that is now exposed to the input pattern that is considered to be the most pertinent. Using a function known as the topological neighborhood function is necessary in order to calculate how the victorious neuron will affect its neighbors. A neighborhood function can be thought of as a representation of a topological map of the region.

$$h_{ji(x)}(t) = \exp \left(-\frac{d_{j,i}^2}{2\sigma^2(t)} \right)$$

4. Results and Discussions

For evaluation, the empirical data is applied over the BC-DRF based food traceability technology. The sensitivity, response time, and accuracy of the approach that is being offered are some of the aspects that are taken into consideration in order to evaluate its effectiveness.

4.1 Experimental Set-up

Python is what is utilized in the process of developing the BC-DRF food traceability system. The computer is built with a random-access memory (RAM) capacity of 8 gigabytes, an Intel i3-core processor, and the Windows 10 operating system as its foundation.

4.2 Dataset Description

Utilizing the food grading information, the BC-DRF food traceability system is evaluated in order to determine its overall efficacy. The dataset that is used to develop machine learning models is built with the help of physical observations. These models may then be used to make predictions regarding the quality of food. The dataset contains seven separate individual metrics, including viscosity, hue, flavor, fragrance, and acidity.

The existence of these components has a considerable bearing on the accuracy with which one can anticipate the quality and grade of food. On a scale that ranges from one (good) to zero (poor), the quality of the food is graded from one (good) to half (moderate), with one being good and half indicating moderate.

Food receives a score of 0 if it does not meet the optimal standards, where the food receives a grade of 1 if it does meet these requirements. If the food does not meet the optimal requirements, then the food receives a score of 0. The temperature as well as the pH are both included in the dataset.

4.3 Comparative Assessment

The recently developed BC-DRF will have its utility evaluated based on how well it compares to the standard procedures used in the industry.

In this section, we examine how the newly created BC-DRF stacks up against traditional methods in terms of a variety of metrics, including accuracy, reaction time, and sensitivity, when applied to a variety of block sizes. The total amount of training data put to use allows for the examination of this comparison to be carried out.

Taking into account a variety of various proportions of the data used for training, Fig. The newly developed BC-DRF-based food traceability system is given a comparative evaluation in figure 8, which uses a block size of 40. The purpose of carrying out this evaluation was to arrive at some decision regarding the efficiency of the procedure.

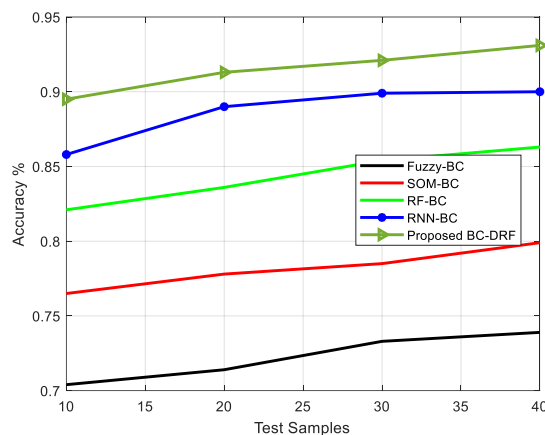


Figure 2: Testing Accuracy

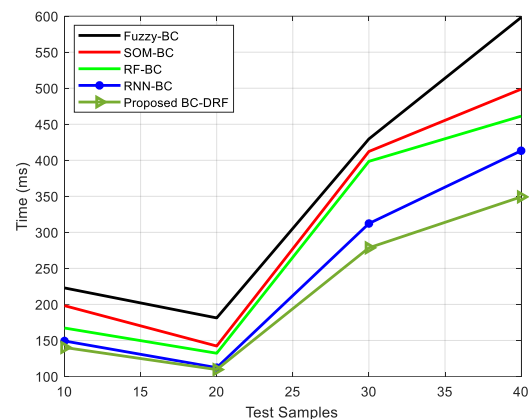


Figure 3: Response time (ms)

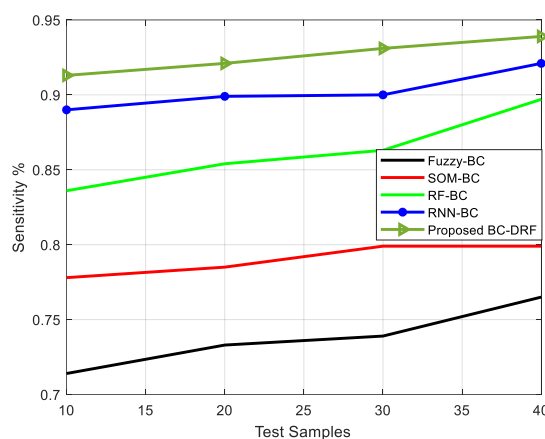


Figure 4: Sensitivity

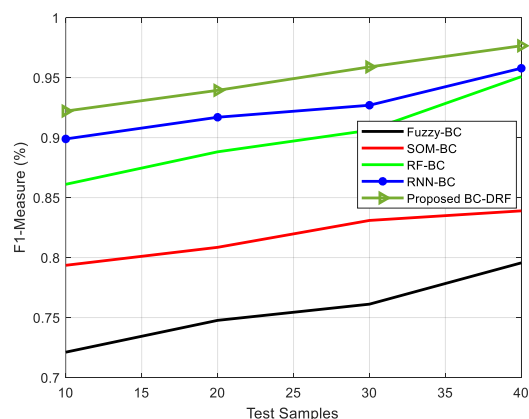


Figure 5: F1-Measure

The findings of an evaluation that were carried out to discover how accurate the newly designed BC-DRF is when it comes to testing are depicted in Figure 8a. This evaluation was carried out to determine how accurate the BC-DRF. The research on response times that was carried out with the help of the adopted BC-DRF methodology is shown in Figure 8(b). Figure 8 (c) displays the results of a sensitivity analysis conducted on the implemented BC-DRF.

A comparison of the newly developed BC-DRF is shown in Figure 2-5, where the values of the metrics correspond to about 90% of the training data.

5. Conclusions

In this paper, we compare the quality evaluation system that is based on BC-DRF with some of the most common existing methods. The system that we use to evaluate the quality of food is based on BC-DRF. When analyzing the differences between the various block sizes, some of the criteria that are taken into consideration include testing precision, sensitivity and reaction time. The use of BC has improved the accuracy of food traceability, while the utilization of

DRF has boosted the efficacy of computing and shortened the reaction time. The utilization of DRF brings about both of these benefits. The fact that the Canberra distance was used in order to extract the features contributed to an additional increase in the system level of sensitivity.

Acknowledgement

This research was not funded by any grant.

References

- [1] Dey, S., Saha, S., Singh, A. K., & McDonald-Maier, K. (2022), "SmartNoshWaste: Using blockchain, Machine Learning, Cloud Computing and QR Code to Reduce Food Waste in Decentralized Web 3.0 Enabled Smart Cities", *Smart Cities*, 5(1), 162-176.
<https://doi.org/10.3390/smartcities5010011>
- [2] Tsoukas. V, Gkogkidis. A, Kampa. A, Spathoulas. G, and Kakarountas. A, (2022). "Enhancing Food Supply Chain Security Through the Use of Blockchain and TinyML", *Information*, 13(5), 213.
<https://doi.org/10.3390/info13050213>
- [3] Mangla. S. K, Kazaçoğlu. Y, Yıldızbaşı. A, Öztürk. C, & Çalık. A. (2022), "A Conceptual Framework for Blockchain-Based Sustainable Supply Chain and Evaluating Implementation Barriers: A Case of the Tea Supply Chain", *Business Strategy and the Environment*.
<https://doi.org/10.1002/bse.3027>
- [4] Sahoo. S, Kumar. S, Sivarajah. U, Lim. W. M., Westland, J. C., & Kumar, A. (2022). "Blockchain for Sustainable Supply Chain Management: Trends and Ways Forward", *Electronic Commerce Research*, 1-56.
<https://doi.org/10.1007/s10660-022-09569-1>
- [5] Chandan, A., John, M., & Potdar, V. (2023). Achieving UN SDGs in Food Supply Chain Using Blockchain Technology. *Sustainability*, 15(3), 2109.
<https://doi.org/10.3390/su15032109>
- [6] Babu, A. S., & Supriya, M. (2022, April), "Blockchain Based Precision Agriculture Model Using Machine Learning Algorithms", In *2022 International Conference on Breakthrough in Heuristics and Reciprocation of Advanced Technologies (BHARAT)* (pp. 127-132). IEEE.
<https://doi.org/10.1109/bharat53139.2022.00036>
- [7] Saba, T., Rehman, A., Haseeb, K., Bahaj, S. A., & Lloret, J. (2023), "Trust-Based Decentralized Blockchain System with Machine Learning Using Internet of Agriculture Things", *Computers and Electrical Engineering*, 108, 108674.
<https://doi.org/10.1016/j.compeleceng.2023.108674>
- [8] Anitha, R., & Rai, D. (2022), "Internet of Things with Artificial Intelligence Detection and Blockchains of Crop Availability for Supply Chain Management", *International Journal of Knowledge-Based Development*, 12(3-4), 444-459.
<http://doi.org/10.1504/ijkbd.2022.128909>
- [9] Khan, A. A., Shaikh, Z. A., Belinskaja, L., Baitenova, L., Vlasova, Y., Gerzelieva, Z., & Barykin, S. (2022), "A Blockchain and Metaheuristic-Enabled Distributed Architecture for Smart Agricultural Analysis and Ledger Preservation Solution: A Collaborative Approach", *Applied Sciences*, 12(3), 1487.
<https://doi.org/10.3390/app12031487>
- [10] Mia, M. U., Rahman, M., Elbeltagi, A., Abdullah-Al-Mahbub, M., Sharma, G., Islam, H. T., & Washakh, R. M. A. (2022), "Sustainable Flood Risk Assessment Using Deep

- Learning-Based Algorithms with A Blockchain Technology”, *Geocarto International*, 1-29.
<https://doi.org/10.1080/10106049.2022.2112982>
- [11] Hannah, S., Deepa, A. J., Chooralil, V. S., BrillySangeetha, S., Yuvaraj, N., Arshath Raja, R., ... & Alene, A. (2022),” Blockchain-Based Deep Learning to Process IoT Data Acquisition in Cognitive Data”, *BioMed Research International*, 2022.
<http://doi.org/10.1155/2022/5038851>
- [12] Neelakandan. S, Rene. B., Prathiba. G, Murthy, L. N, Fantin. I, Arulkumar. N, “Blockchain with Deep Learning-Enabled Secure Healthcare Data Transmission and Diagnostic Model”, *International Journal of Modelling, simulation and Scientific Computing*, 13,04(2022):2472.
<https://doi.org/10.1142/S1793962322410069>
- [13] Hasaan. M, Tayyaba. A, Ahmad. N, Ruzwan. A. N, Woong-Kee. L, “Machines for Classification of COVID-19 Using CT Scans”, *Bioengineering*,10,2 (2023): 203.
<http://doi.org/10.3390/bioengineering10020203>
- [14] Guo. W, Yao.K, “Supply Chain Governance of Agricultural Products under Big Data Platform Based on Blockchain Technology”, *Sci Program*, 2022, 4456150
<http://doi.org/10.1155/2022/4456150>
- [15] Awan. S, “IoT with Blockchain: A Futuristic Approach in Agriculture and Food Supply Chain”. *Wirel Commun Mob Comput.*, 2021, 5580179.
<https://doi.org/10.1155/2021/5580179>

SECUREML BASED CLASSIFICATION FOR INTERNET OF THINGS BASED SECURED TRANSACTION OF DATA

Kesava Rao Alla^{1*}
Gunasekar Thangarasu²

^{1*}Chancellery, MAHSA University, Saujana Putra, Malaysia

²Department of Professional, Industry Driven Education, MAHSA University, Saujana Putra, Malaysia

Abstract: *The Internet of Things (IoTs) is a rapidly developing technology that enables a wide range of applications to interact with each other. The IoT is an emerging technology that can be used to collect, and analyze data from various sources. In this paper, we present SecureML, a unique privacy-preserving CART training scheme that employs blockchain principles to develop a secure CART classifier for use in multipart scenarios in which data is collected from many data sources. SecureML uses the homomorphic cryptosystem to develop secure building blocks such as secure polynomial multiplication and safe comparison. The proposed method is capable of training classifiers in a risk-free manner while retaining an acceptable degree of accuracy.*

Keywords: *Secure, Machine Learning, Classification, Internet of Things, Data Transaction*

1. Introduction

Concerns have been raised regarding the safety and confidentiality of user data, as well as the verification and authentication of devices, as a direct result of the greater interconnection that has taken place [1]. It included approximately two million audio recordings, of which a portion were being held for ransom [2]. With each new day that passes, Internet of Things (IoTs) devices continue to be afflicted by a growing number of security and privacy issues. Preserving both security and privacy in IoT contexts while operating with a restricted number of resources creates significant issues that need to be overcome.

The level of sophistication of attacks is growing at a rate that is directly correlated to the growing difficulty of securing the IoTs. It is possible for desktop PCs and other high-powered computers to identify malware because of the enormous hardware and software capabilities these systems possess. However, there is a shortage of resources for devices that are connected to the IoTs. These more current types of cyber protection do not require regular upgrades in the same way as traditional cybersecurity systems and software are unable to notice minor attack variants or zero-day attacks [3]. In other words, these more recent types of cyber protection do not require regular updates. Another one of the service provider shortcomings is that it does not provide real-time updates, which leaves the network open to attack [4].

Not only may the implementation of machine learning (ML) technologies be useful for cybersecurity systems, but they can also be useful for the infrastructure of the IoTs. These algorithms could maintain system security against new attacks while also updating threat knowledge databases, evaluate network traffic based on what is already known about cyberthreats, and evaluate network traffic based on what is already known about cyberthreats [5].

In addition to the strategies based on machine learning, researchers have begun employing the cutting-edge Blockchain (BC) technology in order to safeguard the underlying systems [6]. The prevalence of cyber dangers in the IoT business has resulted in the development of both ML algorithms and BC methodologies. Nevertheless, the combination of the two is novel, and it should be the subject of more research [7].

Incorporating BC into the training process for ML holds potential but is not without its share of obstacles. The first challenge that has to be addressed is conceiving of a data format for the training that is easily supported by blockchain technology while preserving the confidentiality of each unique service provider at the same time [8]. This is the first issue that needs to be addressed. The second obstacle is to devise a training algorithm that, without disclosing any sensitive information, uses BC data to build reliable classifiers. This is a particularly difficult task [9] [10], [15]-[18].

SecureML is a privacy-protecting CART Classifier based training system that makes use of blockchain technology to encrypt data from IoTs devices. SecureML uses a public-key cryptosystem, it is necessary for data providers to encrypt their data locally using their own private keys. This is a requirement of the secureML standard. This contributes to ensuring that the data collected by IoT devices is kept confidential.

2. Related Works

Reyna et al. [11] study how BC approaches might have the potential to enhance the data reliability and security of the IoTs. According to the findings of the research that was carried out, one of the potential problems that can be caused by adopting BC techniques is an increased risk of having one security violated. During the course of the investigation, a number of potential vulnerabilities in the system security were uncovered. In addition, the study highlighted the several ways in which business continuity strategies, applications, and platforms can be utilized in conjunction with the IoTs. The fact that this study did not evaluate a wide variety of additional security threats associated with the IoTs was one of the poll major flaws. The survey also had a number of other limitations.

After separating the various layers of security that make up the IoTs into more manageable pieces, Khan et al. [12] disclosed their BC-based solutions. It has been discovered that there are low-level, middle-level, and high-level security concerns associated with the IoTs. Insecure interfaces, access to the internet, software vulnerabilities, and middleware security are among the concerns of substantial magnitude. In addition, the research presented alternative solutions for each problem. The research performed an excellent job of bringing to light the threats that each layer of the IoTs is exposed to; however, it did not provide any specifics of the ways in which BC can be used to protect against these threats.

In addition to analyzing the security problems and new developments in integrating IoT with blockchain technology, Panarello et al. [13] provided a complete review of the BC consensus procedures. In addition to this, the authors talked about integrating the IoTs. We carried out a comprehensive investigation of the most recent BC-based solutions, and the accumulated material was arranged in accordance with the application fields it covered.

Wang et al. [14] outlined the various ways in which the IoTs may be hacked, such as through its endpoints, communication lines, network protocols, data acquired by sensors, distributed denial-of-service attacks, and malicious software. One of the ways in which the IoTs could be hacked is through the data collected by sensors. The development of a framework for the IoTs that is based on BC may offer a variety of potential advantages and motivations. The outcomes of the research indicate that a centralized IoTs paradigm is vulnerable to Distributed Denial of Service attacks. Because the architecture is dependent on a single point of failure, there is a greater possibility that services related to the IoTs will not be accessible when they are required.

Hassija et al. [15] analyzed the parallels and differences between IT security and IoT security. Following this, they grouped IoT applications in accordance with the security and privacy challenges that they raise. This research went farther and investigated potential security vulnerabilities across all four layers of software used in IoTs devices. This study is possibly the most analogous to the work that we have done; nevertheless, it focuses on security and does not address the problems regarding privacy that are specific to the IoTs..

3. Proposed Method

In this research, we introduce SecureML, which is a unique privacy-preserving CART training scheme that employs blockchain principles to develop a secure CART training algorithm for use in multi-part scenarios in which IoT data is collected from many data sources. This approach is intended for use in situations where numerous data providers contribute IoT data. This solves the problems of both protecting the privacy of user data and maintaining the accuracy of that data, both of which are significant hurdles in the current technological world.

3.1 System Model

The research plan is to create an ecosystem for the IoT that is powered by data with the help of IoT devices, data suppliers, data analyzers, and an IoT platform that is based on blockchain technology.

- IoT devices are able to collect and send vital information via either wireless or wired connections, regardless of the type of connection they are using. The information that is obtained in today smart cities may be put to use in a variety of different ways, and this includes anything from weather reports to biometric data. Among the information that can be gathered are biometric measurements. Because of their limited processing capabilities, IoT devices will not be able to contribute to the procedures that are being carried out involving the sharing and analysis of data.
- The data providers for IoTs devices are the ones who are responsible for gathering the information that is generated by those devices. Sensitive information is routinely included in data sets pertaining to the IoTs due to the fact that it is such a valuable asset to data providers. Therefore, in order to preserve the IoTs data that it has collected, each data source encrypts it using a partly homomorphic key before adding the information to the distributed ledger.
- The platform for the IoTs, which is built on blockchain technology, acts as a distributed database by recording encrypted data from the IoTs that is gathered from all data sources. A system that is embedded within IoT devices ensures that the data that is exchanged between those devices is correct and comprehensive.

If they employ cutting-edge tools, IoT data analysts are optimistic that they will be able to acquire a comprehensive understanding of the IoT data that is stored on blockchain-based platforms. This is because the analysts have great hopes that they will be able to do so. In order to acquire the training CART classifier parameters, analysts are going to need to collaborate with the various data sources that are applicable.

3.2 Threat Model

According to the model of the system that is shown in Figure 1, there are many different kinds of threats that can entities, and there are also many different kinds of interactions that can take place between those entities. We set out to develop a method that respects the privacy of users while simultaneously training CART models on data from a range of IoT sources. To do this, we focused solely on the interactions that occur between data suppliers and data analysts when developing this method.

The data analyst is viewed as an enemy who is not just trustworthy but also interested in the matter at hand. By sticking to the predesigned ML training protocols, the data analyst would be behaving in good faith; nevertheless, it is possible that he would get drawn in by the contents of the data and attempt to learn more by decrypting and analyzing the encrypted data as well as the intermediate data of processing. In this case, he would not be acting in good faith. The fact that the data analyst has access to confidential information serves as a guiding principle for our examination of two threat models that are extensively utilized and have strikingly dissimilar attack capabilities.

- **Known Ciphertext Model:** It is stored on a blockchain-based platform, which enables encryption, only the IoT data analyst has access to the encrypted IoT data. This keeps the data secure and prevents unauthorized access. The IoT data analyst is also able to monitor the iteration step and descent gradient that are created by the operation of the secure training algorithm. This is an additional capability of the IoT data analyst.
- **Known Background Model:** It is generally accepted that the IoT data analyst is in possession of a higher breadth of knowledge in the known background model than is attainable in the known ciphertext model. An IoT data analyst may be able to receive access to the sensitive information stored by a separate IoT data provided by collaborating with one or more other IoT data providers. This may allow the analyst to obtain access to the confidential information.

We make it possible for any three or more IoTs data suppliers as well as any IoTs data analyzer to collude together in order to violate the privacy of participants by providing the means for them to do so. For the sake of clarity, let presume the following: Each player adopts the role of a trustworthy opponent who abides by the rules of the protocol but who may also be interested in the confidential information of other domains. This adversary plays by the rules of the protocol. Any two or more individuals are free to work together, and participation in collaborative efforts is actively encouraged.

They are adversaries that follow the protocol, but based on the values they acquire, they draw conclusions about the privacy of other people. The characteristic of these enemies is that they are passive. Because of the method that we have developed, any private information that is gathered for the CART model training will be kept strictly confidential. To give more clarity,

the privacy of the IoTs is composed of the records of the IoTs that belong to each data provider as well as the CART model parameters that belong to each data analyst. The following is a rundown of the various levels of safety that we intend to achieve:

- 1) The privacy of the IoT data analyst, as well as the privacy of each IoT data provider, is protected against dishonest opponents who are nonetheless curious. This protection extends to the privacy of the IoT data itself.
- 2) The privacy of the IoT data analyst and the privacy of each IoT data provider shall be safeguarded in the event that two or more parties are involved in a conspiracy with one another. This protection will apply even if the conspiracy involves more than two parties.

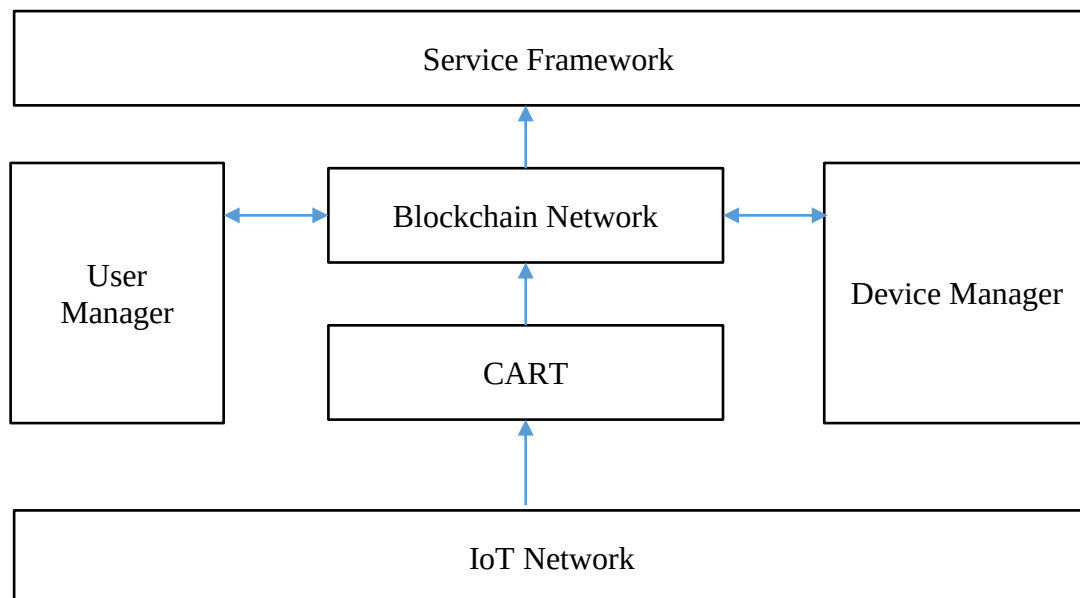


Figure 1: Proposed Framework

Using the encrypted data that is kept in the global ledger, a data analyst can develop a training method in order to train a CART model. This algorithm is capable of making use of a number of different components, some of which include safe comparison and safe polynomial addition. During the course of the training, it will be necessary to convey intermediate outcomes, which will demand interactions between the data analyst and each data source

3.3 SecureML Model

The goal of this approach to generating classification trees is to pick each subset in such a way that the newly produced subgroups are purer than the initial set of data. It is planned to develop subsets of the data that, in regard to the variable that is being researched, are as comparable to one another as is practically achievable. This goal is intended to be achieved as soon as possible.

When different statistical units are compared to one another, the presence of impurities can be illustrated by differences in the means of the dependent variable. Homogeneity, also known as purity, is achieved when all n statistical units share the same mode. This is the case when there

are n statistical units. This is based on the assumption that the r -mode variation of the qualitative phenomena in question is being discussed.

The level of heterogeneity will be at its highest when the statistical units are evenly distributed throughout all r modes in such a way that the relative frequency of each mode is equal to $1/r$. This will result in the highest level of heterogeneity possible. Because the principle of decomposition in groups and between groups also applies to heterogeneity, the parent node two offspring nodes have a higher degree of similarity to one another than the parent node does. This is due to the fact that the property of decomposition in groups as well as between groups is applicable to heterogeneity as well. Despite the fact that it offers a number of benefits, the CART method can only be applied to a binary classification scheme.

The following function calculates the degree of contamination that is existing at node t in the network.

$$i(t) = \Phi[f_1, f_2, \dots, f_j]$$

where $\Phi(.)$ - nonnegative function

A node impurity is at its highest when it contains cases from all classes of the dependent variable, and it is at its lowest when it contains cases from only one class if $(.)$ is a positive function. If this is the case, then a node impurity is at its largest when it contains cases from all classes of the dependent variable. If $(.)$ is a positive function, then for the purpose of our research, we utilized the Gini heterogeneity index to determine the degrees of impureness present in the data.

$$i(t) = 1 - \sum_{j=1}^J f_j^2$$

In the most extreme instance of homogeneity, it has a value of 0 (also known as zero heterogeneity), whereas in the most extreme occurrence of heterogeneity, it has a value of $(r - 1)/r$.

The formula that is used to define the amount that the impurity of the node t decreases as a direct result of each given split (s):

$$\Delta i(s, t) = i(t) - f_l \times i(t_l) - f_r \times i(t_r)$$

where

f_l - cases proportion in a left node and

f_r - cases proportion in a right node.

$\Delta i(s, t)$ - non-negative value.

When building the classification trees, for any given node t , the split s^* that results in the greatest possible decrease of impurity in the tree is selected as the one to use in the construction process. This is achieved by creating all conceivable dichotomizations of the explanatory variables in a way that is consistent with the properties that are intrinsic to them.

$$\Delta i(s^*, t) = \max_{s \in \Phi} \Delta i(s, t)$$

where

Φ - subdivisions set of a IoT node t .

The structure of the tree s^* is designed in such a way that every node and branch has its own peak. It has been proved that selecting the split that results in the greatest decrease in impurities $\Delta imp(s, t)$ is equivalent to selecting the split that results in a lower overall impurity level in the shaft. This was done by comparing the two split final impurity levels. the local optimization criterion of a classification tree and the global optimization criterion of the tree are precisely the same.

Two constraints were imposed in order to stop the tree from expanding any further: (a) the impurity reduction threshold was set at 0.001, and (b) the maximum size of the tree was set at five levels. Both of these requirements were satisfied. We decided to use a posterior classification ratio (PCR) to assign response class to each node of the tree rather than the median as our measure of central tendency because we wanted to figure out the factors that are responsible for the observed shift in response.

This was done in an effort to establish the factors that are responsible for the observed shift in LSS response. This was done in an effort to separate the factors that are responsible for the observed shift so that we could examine them individually. In order to calculate the PCR, a number of processes have to be done, including the following:

$$PCR(j) = p(j)/p(j|root)$$

where t_{root} - root node.

In the made-up scenario in which the posterior classification ratio was exactly 1, the evidence acquired from the posterior distribution would be split evenly between the two classes. This would be the case since the ratio would be exactly 1. There is no indication that one group is more likely than another when looking at the data as well as the previous distributions.

For a positive classification, the likelihood of a positive posterior distribution increases if the value is larger than 1.0; on the other hand, for a negative classification, the likelihood of a positive posterior distribution increases if the value is less than 1.0. It was decided to give preference to the j^* class that had the higher PCR value while classifying each node, which brought about the categorization of each node.

4. Results and Discussions

An intelligent fitness blockchain network, fitness IoTs servers, IoT gateway implementations, fitness blockchain front-end development, and a predictive analytics model are the five components that make up the overall execution of the suggested system. The proposed system underwent extensive testing on a computer of the eighth generation equipped with an Intel Core i-5 CPU, 8 GB of RAM.

According to our design, it is the job of each IoT data provider to first collect all of the data that is generated by IoT devices that are located within its specific domain and then process that data in the manners that are described below. This responsibility begins with the collection of all of the data that is generated by IoT devices that are placed within its specific domain. This device also functions as an IoTs data analyzer in addition to its other duties as a data provider.

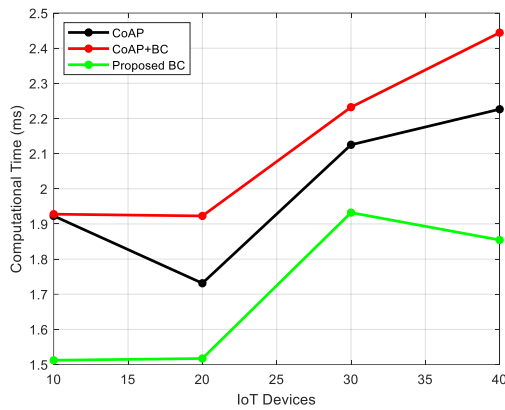


Figure (a): Testing

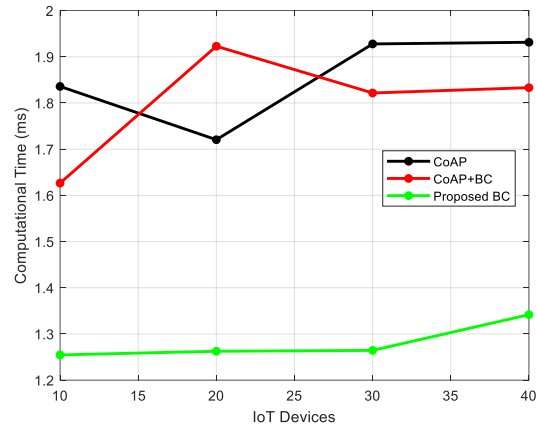


Figure (b): Training

Figure 2: Computational Time

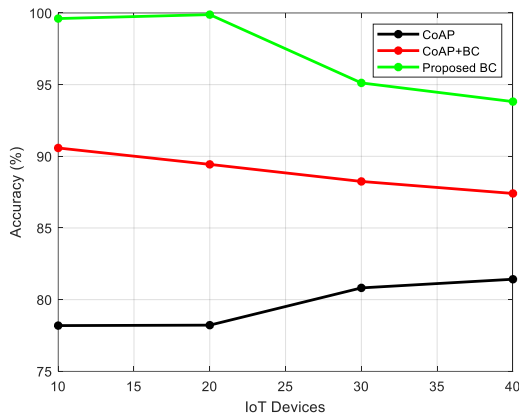


Figure (a): Testing

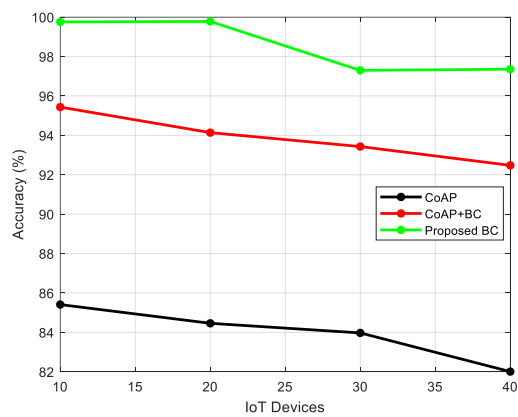


Figure (b): Training

Figure 3: Accuracy

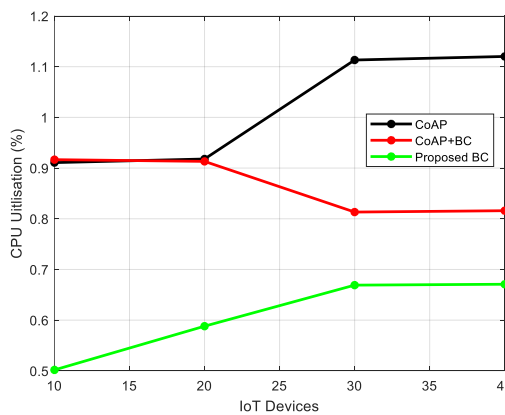


Figure (a): Testing

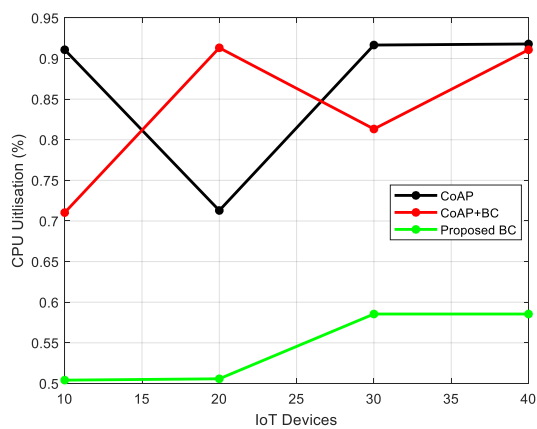


Figure (b): Training

Figure 4: Response Time

We constructed the CART to demonstrate that secureML does not compromise accuracy while simultaneously preserving the privacy of any IoT data provider and safely training classifiers. Because of this, we were able to show that using secureML did not result in a loss of accuracy. We will not be making any adjustments to the training parameters as part of the work that we are undertaking; rather, the focus of our efforts will be on determining how to train a classifier in a secure manner.

Another experiment was changing the total number of endorser peers who are connected to the blockchain. In the context of a possible commercial exchange, the peer node whose recommendation is being sought is referred to as an endorser peer. Just before to a transaction being permanently committed, the endorser-peer function of the blockchain.

Figure 2 shows how the computational latency is affected when there is a shift in the peer node transmission rate between the ranges of 25 and 200 TPS. It has been shown that the amount of time it takes for the delay to occur increases in a manner that is directly proportional to the number of peers that are participating in the network.

Figure 3 depicts how well the performed after we changed the total number of peer nodes and then measured the accuracy. The amount of time it takes for data to travel between peer nodes in the network grows in direct proportion to the number of peer nodes that are there. The network peer node creates a bigger demand on the system, the throughput of the node is slowed down as a result.

Figure 4 shows that the CPU utilization throughput of a network drops as the number of peer nodes in the network increases. This is something that can be observed. This is due to the fact that the existence of a greater number of nodes causes an increase in the volume of traffic.

5. Conclusions

In order to ensure the integrity of data transfer between the many different IoTs devices, we make use of blockchain technology. The data instances that are provided by each IoT data provider can be encrypted locally using a private key, and this encrypted data can then be deposited on the blockchain using transactions that comply to a predefined transactional protocol. In other words, the IoT data providers can encrypt the data instances that they offer to each other. We are able to establish a safe and reliable CART training algorithm with the assistance of the homomorphic cryptosystem by developing secure building blocks such as secure polynomial multiplication and safe comparison. This strategy calls for only two exchanges throughout each iteration and eliminates the need for a dependable third party as a prerequisite. Extensive testing is done to demonstrate that our method is capable of training CART classifiers in a risk-free manner while retaining an acceptable degree of accuracy. This is done in order to show that our method is superior to other approaches.

Acknowledgement

This research was not funded by any grant.

References

- [1] Ogundokun, R. O., Arowolo, M. O., Misra, S., & Awotunde, J. B., "Machine Learning, IoT, And Blockchain Integration for Improving Process Management Application

- Security”. In *Blockchain Applications in the Smart Era*, Cham, Springer International Publishing, (2022) :237-252.
[https://doi.org/ 10.1007/978-3-030-89546-4_12](https://doi.org/10.1007/978-3-030-89546-4_12)
- [2] Hu, H., Xu, J., Liu, M., & Lim, M. K. “Vaccine Supply Chain Management: An Intelligent System Utilizing Blockchain, IoT and Machine Learning”. *Journal of Business Research*, 156 (2022) :113480.
<https://doi.org/10.1007/s11277-020-07802-4>
- [3] Mozumder, M. A. I., Sheeraz, M. M., Athar, A., Aich, S., & Kim, H. C. (2022). “Overview: Technology Roadmap of the Future Trend of Metaverse Based On IoT, Blockchain, AI Technique, And Medical Domain Metaverse Activity”. In *2022 24th International Conference on Advanced Communication Technology*, (2022): 256-261.
<https://doi.org/10.1155/2019/3174304>
- [4] Javed, A. R., Hassan, M. A., Shahzad, F., Ahmed, W., Singh, S., Baker, T., & Gadekallu, T. R. (2022), “Integration of Blockchain Technology and Federated Learning in Vehicular (IoT) Networks: A Comprehensive Survey”, *Sensors*, 22, no.12 (2022): 4394.
<https://doi.org/10.1109/JIOT.2021.3065691>
- [5] Al-Qarafi, A., Alrowais, F., S. Alotaibi, S., Nemri, N., Al-Wesabi, F. N., Al Duhayyim, M, & Al-Shabi, M. “Optimal Machine Learning Based Privacy Preserving Blockchain Assisted Internet of Things with Smart Cities Environment”, *Applied Sciences*, 12, no. 12 (2022) :5893.
<https://doi.org/10.1049/iet-com.2018.5109>
- [6] Vairagade, R. S., & SH, B., “Enabling Machine Learning-Based Side-Chaining For Improving Qos In Blockchain-Powered IoT Networks”, *Transactions on Emerging Telecommunications Technologies*, 33, no.4 (2022): 33-44.
<https://doi.org/10.1109/JIOT.2019.2898166>
- [7] Mrabet, H., Alhomoud, A., Jemai, A., & Trentesaux, D., “A Secured Industrial Internet-of-Things Architecture Based on Blockchain Technology and Machine Learning for Sensor Access Control Systems in Smart Manufacturing”, *Applied Sciences*, 12, no.9 (2022): 4641.
<https://doi.org/10.1016/j.jksuci.2018.04.004>
- [8] Whig, P., Velu, A., & Nadikattu, R. R., “Blockchain Platform to Resolve Security Issues in IoT and Smart Network”, In *AI-Enabled Agile Internet of Things for Sustainable FinTech Ecosystems*, IGI Global, (2022): 46-65.
<http://doi.org/10.1109/JSEN.2021.3055948>
- [9] Wu, Y., Jin, X., Yang, H., Tu, L., Ye, Y., & Li, S., “Blockchain-Based Internet of Things: Machine Learning the Sensing Trusted Traceability System”, *Journal of Sensors*, (2022)
<https://doi.org/10.1016/j.jpdc.2020.04.007>
- [10] Ahmed, I., Zhang, Y., Jeon, G., Lin, W., Khosravi, M. R., and Qi, L., “A blockchain-and Artificial Intelligence-Enabled Smart IoT Framework for Sustainable City”, *International Journal of Intelligent Systems*, 37, no. 9 (2022): 6493-6507.
<https://doi.org/10.1016/j.comcom.2019.03.007>
- [11] Reyna, A., Martín, C., Chen, J., Soler, E., and Díaz, M., “On Blockchain and Its Integration with IoT. Challenges and Opportunities”, *Future Generation Computer Systems*, 88, (2018): 173-190.
<http://doi.org/10.1109/SCORED.2019.8896238>
- [12] Khan, S., Khan MA, Salah K, “IoT Security: Review, Blockchain Solutions, and Open Challenges”, *Future Generation Computer System*, 82, (2018):395-411.
<https://doi.org/10.3390/sym14122472>

- [13] Panarello, A., Tapas, N., Merlino, G., Longo, F., and Puliafito, A., “Blockchain and IoT Integration: A Systematic Survey”, *Sensors*, 18, no. 8, (2020):2575.
<http://doi.org/10.1109/ICCI51257.2020.9247795>
- [14] Wang, X., Zha, X., Ni, W., Liu, R. P., Guo, Y. J., Niu, X., & Zheng, K., “Survey on blockchain for Internet of Things”, *Computer Communications*, 136, (2019):10-29.
<https://doi.org/10.1007/s00500-021-06462-1>
- [15] Hassija, V., Chamola, V., Saxena, V., Jain, D., Goyal, P., and Sikdar, B., “A survey on IoT security: application areas, security threats, and solution architectures”, *IEEE Access*, 7, (2019); 82721-82743
<http://doi.org/10.1109/ACCESS.2019.2924045>
- [16] Hannah, S., Deepa, A. J., Chooralil, V. S., BrillySangeetha, S., Yuvaraj, N., Arshath Raja, R, and Alene, A., “Blockchain-based deep learning to process IoT data acquisition in cognitive data”, *BioMed Research International*, 2022.
<https://doi.org/10.1155/2022/5038851>
- [17] Singh.V., and Dahiya. N., “Blockchain-Based Federated Machine Learning for Solving IoT Security Problems”, In *Applications of Blockchain and Big Iot Systems*, Apple Academic Press. (2022): 27-56.
<https://doi.org/10.1155/2022/5038851>
- [18] Dorri, A., Roulin, C., Pal, S., Baalbaki, S., Jurdak, R., and Kanhere, S. S., “Device Identification in Blockchain-Based Internet of Things”, *IEEE Internet of Things Journal*, 9, no. 24, (2022); 24767-24776.
<https://doi.org/10.1109/JIOT.2022.3194671>

SOFTWARE MODEL CHECKING DISTRIBUTED APPLICATIONS: A HYBRID APPROACH

Hing Ratana¹
Sharifah Mashita Syed-Mohamad^{2*}
Chan Huah Yong¹

¹School of Computer Sciences, Universiti Sains Malaysia, 11800 Penang, Malaysia

² Faculty of Ocean Engineering Technology and Informatics, Universiti Malaysia Terengganu, 21030 Kuala Nerus, Terengganu, Malaysia

Abstract: *Developing reliable distributed systems poses significant challenges due to the non-deterministic execution of threads, processes, and communication channels. Formal verification methods, such as model checking, play a crucial role in ensuring the reliability of safety-critical systems. Model checking explores the complete behavior of the system under test (SUT) systematically, investigating each reachable state with different thread schedules. Recent software model-checking techniques, namely cache and centralization, have been applied to distributed systems. However, the caching technique can only check one process at a time, while the centralization technique verifies all processes simultaneously. In the centralization technique, two "ArrayByteQueue" buffers are utilized to store communication and process data byte-by-byte. However, during the model-checking process, the read and write operations, involving data extraction and removal from the queue, become resource-intensive due to the backtracking process. As a consequence, existing interprocess communication (IPC) models encounter computational limitations and experience a rapid state space explosion. To address these challenges, our work proposes the remodeling of IPC models by introducing a request and response tree structure to store communication data. Additionally, we employ pointers to navigate through the data during the backtracking process. Through experimental evaluations, the proposed implementation choices have demonstrated significantly improved performance across various metrics. By incorporating the request and response tree, we enhance the efficiency of storing communication data, while the use of pointers optimizes navigation during backtracking. This remodeling of IPC models shows promise in mitigating computational limitations and state space explosion, thereby enhancing the model-checking process in distributed systems. Our research contributes to advancing the field of model checking in distributed systems and offers potential solutions to the challenges associated with resource-intensive read and write operations during the model-checking process*

Keywords: *Software Model Checking, Distributed System, Java PathFinder (JPF), Interprocess Communication (IPC), Backtracking, State Space Explosion.*

1. Introduction

Modern societies nowadays rely on software in all aspects of daily living and interacting with each other, including safety-critical systems, electric cars, smart home devices, and so on. Software or the applications such as Google Search, Facebook, YouTube, Netflix, Shopee, Lazada, Grab, Foodpanda, and so many more are distributed. A distributed system consists of a group of nodes. Those nodes can be generically referred to as physical smartphones, personal

computers, smart devices, televisions, and so on which cooperate by exchanging messages over communication links to achieve some tasks [1].

This article focuses on software model checking of distributed systems. Catching bugs earlier as possible. Big companies like Google have spent most of their cost focusing on finding and fixing, processes, and procedures to debug incidents in Google's distributed systems [2]. Many large software systems consist of multiple processes spread across multiple computer processors, executing independently. While this provides the advantages of increased performance and scalability it also makes such systems much harder to test due to partial failure and asynchrony. Partial failure refers to the components in distributed applications that can fail along the way, resulting in incomplete results or data. Asynchrony is the indeterminateness of ordering and timing within a distributed system that often leads to solutions with a high degree of complexity. Avoiding distributed system bugs also requires reasoning about the integration between nodes and must tolerate the failure of the underlying hardware. In addition, the probability of human error in either design, implementation, or operation also contributes to system bugs. Therefore, developing a reliable distributed system is a very challenging task.

Formal methods, particularly model-checking, can produce rigorous and automated reliability proofs for hardware and software systems. The area has focused on distributed systems for two main reasons [3]. First, distributed programs are error prone, because programmers must consider all possible effects induced by different scheduling of events. Second, testing, which is widely used for certifying sequential programs, tends to have low coverage in distributed settings, because bugs are usually difficult to reproduce. They may happen under very specific thread schedules, and the likelihood of taking such corner-case schedules may be very low. Therefore, automated verification techniques represent crucial support in the development of reliable distributed applications.

Model checking [4] is a technique to detect property violations in a concurrent system by exploring every possible execution path. Accordingly, every possible state of the system is checked against given properties. This technique is very useful for the quality assurance of safety-critical systems and core algorithms/protocols of large systems. Model checking is originally developed for hardware verification, but the concept of state space exploration has been applied to a wide range of software systems as well.

In the traditional software model-checking process, a system to be verified is abstracted into an input language supported by the model checker, a software tool that performs model-checking [5]. The model checker then generates a directed graph that represents the state space of the system. It traverses the graph and checks if the desired properties hold at every state. This activity is referred to as state space exploration. After verification of the model, the system is usually implemented in a programming language such as C or Java, based on the verified model.

Although model checking originally requires the system to be checked written in formal languages, e.g., PROMELA, recent work in the community applies model checking directly to source codes; this activity is often called software model checking. Direct verification of real codes increases confidence in software safety. The fact that the system design meets a specification does not imply that the implementation does. Many concurrency-related bugs are still introduced by programming mistakes during the implementation phase, such as race

conditions, deadlocks, and assert violations. Verification in the design phase cannot guarantee the absence of bugs in the final deliveries. The software model-checking community focuses on the analysis of real implementations, written in mainstream programming languages such as Java or C [6].

This article focuses on utilizing software model checking in the context of distributed systems using Java. The research aims to enhance existing techniques by combining the cache approach with centralization. The centralization technique [7] verifies all processes within the distributed systems simultaneously. This technique employs two 'ArrayByteBuffer' buffers to store communication and process data at a byte-by-byte level. However, the write and read operations, which involve extracting data from the queue during the model-checking process, pose significant resource-intensive challenges due to the backtracking process. This leads to a rapid expansion of the state space and computational limitations in existing interprocess communication (IPC) models. Therefore, our study proposes a novel remodeling of the existing IPC models by incorporating a request and response tree for storing communication data and employing pointers to navigate through the data during the backtracking process. The paper is organized into the following sections: Section 2 provides an overview of recent advancements in model-checking distributed systems, including caching and centralization techniques. Section 3 presents our methodology and the proposed solution. Section 4 offers insights into the experiments conducted and the results obtained from the proposed method. Finally, in Section 5, we conclude our work.

2. Literature Review

Initially, model checking was developed as a means to verify algorithms, protocols, and system models. However, its scope has expanded to include direct verification of software systems or implementations. With the increasing complexity of code compared to models, there has been a growing interest in model-checking networked applications. This paper focuses on explaining the techniques relevant to model-checking networked applications.

Model checking involves exhaustively and systematically examining the behavior of a system under test (SUT) by analyzing each reachable state for different thread schedules. The SUT undergoes backtracking during the model-checking process. In the context of model-checking network applications, the SUT often engages in repeated messaging (input/output operations) with external processes. However, these external processes, which are not controlled by the model checker, lack coordination with the SUT's backtracking mechanism, resulting in a breakdown of direct communication between the SUT and the external processes.

Among model checkers, Java PathFinder [8] stands out as the only one capable of verifying distributed systems, as demonstrated in the 2019 software verification competition contribution [9]. Various approaches, such as cache [10]–[12] and centralization [7], [13]–[17], have been proposed to tackle the challenges of model-checking distributed systems.

2.1 Cache Approach

The cache approach model checks a single process inside the model checker tool once at a time and runs all the other processes externally in their native environment. A process is a self-contained execution environment and has its resources such as memory and other resources, whereas threads run within a process and share the process runtime resources. In the cache-based approach, the SUT and peers denote the single process inside the model checker and the

external processes, respectively. The SUT executed by the model checker is subjected to backtracking, while external processes run normally.

Figure 1 illustrates the overall architecture of the cache approach for model-checking network applications. The model checker executes the single-process SUT in exhaustive ways thus making repeated requests. The special cache layer intercepts all the communications between the SUT and its peers. It represents the state of communication at different points in time. Data previously received by the SUT is held in the cache and if the SUT makes the same request, the response will be sent from the cache rather than incur all the processing involved in resending the request to the peers.

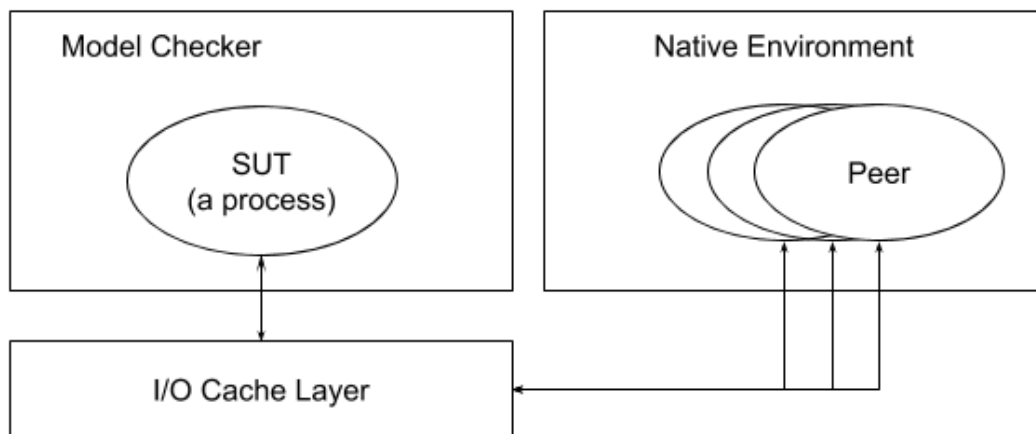


Figure 1: Overall Architecture of Cache Approach

The main challenge of this approach is the synchronization between the single-process SUT and its peers since the SUT is subjected to backtracking by the model checker, and the model checker does not have any control over its peers. During model checking SUT, the SUT may resend data which might interrupt the correct behavior of the peers, and the peers may not send the correct data back to the SUT. A special cache layer has been developed to solve these problems. Existing cache-based techniques address this problem by introducing a special cache layer between the SUT and its peers for state synchronization.

Initial work by [10] proposes a solution for model-checking network applications by developing a special caching layer for stream-based input/output (I/O). They introduce the idea of I/O caching via deterministic communication. The method is called linear-time cache. A linear cache models communication data using a pair of arrays called a request array and a response array. The linear data structure stores request and response messages, respectively. The initial solution works if the I/O operations of the SUT always produce the same data stream regardless of the indeterminacy of the thread schedules. The communication between the SUT and its environment must be independent of the thread schedules. For instance, if the client sends a sequence of characters to the server, the server supposes to send the same sequence of characters back to the client, regardless of the thread schedules. If it is not the case, the behavior of the communication traces would be undefined.

The later work by [11] extends the idea of caching I/O traces to a wider range of network applications by developing the tree data structure that allows diverging communication traces between different thread schedules. This concept is called branching-time cache. The technique

captures the communication traces and stores them in a tree data structure. The major advantage of branching cache is to allow the non-determinism of thread schedules within the SUT, but it does not allow non-determinism within the peers. For this approach, the SUT at least can send sufficient different data from the previously observed ones.

To allow non-determinism within peers, the proposed work by [11] combines branching-time cache with process checkpointing [12]. This is called the hybrid approach. Process checkpointing environment can run, pause, and replay the peers at any point in time. During model checking of SUT, the checkpointing idea can be incorporated when the SUT requires the synchronization of data from the peers, at those points, checkpointing can play and replay the peers' states accordingly to the requests from the SUT. By doing this, this concept gives a broader range of model-checking network applications.

2.2 Centralization Approach

The centralization approach is to model-check all processes within a model checker. These techniques can be applied at SUT Level, OS Level, and model checker level. Figure 2 shows the overall architecture of existing centralization techniques: (a) SUT level; (b) OS level; (c) Model checker level. At the SUT level, the processes are transformed into one main process, so each process is mapped into a thread, and the model checker verifies the main process. At OS-level centralization, it does not involve transforming the SUT. Instead, all the processes are running on top of the virtualization tool, and the model checker tool is extended to capture the state of the virtualization tool for state-space exploration. Finally, the centralization approach can be applied at the model checker level. This technique is to extend the model checker so that the tool can capture multiple processes within the tool itself.

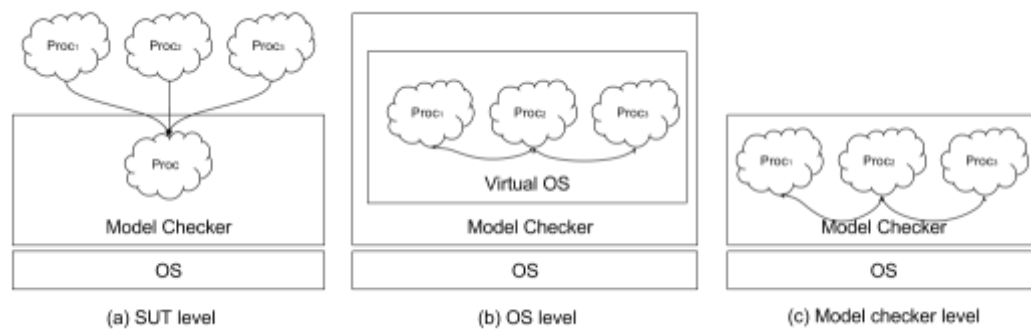


Figure 2: Overall Architecture of Centralization Approach

2.2.1 SUT Level Centralization

Initial work by [13] applying the centralization technique at the SUT level. They propose the concept of transforming processes into a single process by replacing remote method invocation (RMIs) with local ones that simulate RMIs. In addition, Stoller and Liu develop CentralizedThread class that extends Thread and initialize an instance of field type integer to denote process id. By doing this, they can map each process into a thread, and each thread communicates with the other via the simulated local RMIs. Stoller and Liu initially introduced centralization at the system level, which involves merging all distributed Java application processes into one and replacing remote method invocations with simulated local ones. Java RMI, which enables the communication between Java processes, is used for objects in different JVMs to call each other's methods remotely.

Later work by [14] provides a more accurate transformation of processes into a single process, and they also address some of the limitations of previous work by Stoller and Liu. In contrast to previous work, Artho and Garoche perform bytecode instrumentation which applies to systems compatible with a newer version of Java, and, in addition, their technique is also applicable to applications that use sockets for communication.

Like the solution proposed by Stoller and Liu, the approach presented by Artho and Garoche replaces calls to methods that end the process, such as "System.exit(int)", by throwing the "java.lang.ThreadDeath" exception. However, their approach also doesn't offer a mechanism for killing other threads within the terminating process. If all remaining threads are daemons, their solution suggests ignoring the failures in a dead process. They do provide a mechanism for freeing resources after processes terminate, but starting shutdown hooks requires manual modification of the centralized program. On the other hand, the RMI model proposed by Stoller and Liu cannot be extended to support communication via sockets, while the centralization technique of Artho and Garoche applies to applications that communicate through sockets.

also, adopt centralization at the SUT level, and their approach builds upon the work of Artho and Garoche while addressing some of its limitations. Specifically, they address the class version conflict issue, which arises when different processes use classes with the same name but different bytecode. Unlike other proposed methods, their approach doesn't require processes to use the same version of Java classes. Before implementing their centralization algorithm, they resolve class version conflicts between processes by renaming classes that have the same name but different bytecode. However, their approach does not resolve class version conflicts for Java system libraries, as it does not apply to native methods and code that utilize the reflection API.

Ma et al's [15] approach provides a method to end processes by ending all their threads, which they accomplish by using the Java thread interruption mechanism. However, this requires adding code to the processes. Additionally, their approach does not handle starting shutdown hooks when a process terminates.

Finally, SUT-level centralization has been proposed by [16]. They are mainly focusing on environment generation by introducing a framework called Netstub. The Netstub API requires users to manually develop how the environment should be generated to accommodate the SUT during model checking. In addition, Netstub also allows model checking a process at a time. The Netstub environment can generate network events that are perceived by the SUT.

2.2.2 OS Level Centralization

In centralization at the OS level, the processes are running on a virtualization tool; therefore, this approach does not require transforming the SUT. This approach requires the extension of the model checker's scope to capture the state of the virtualization tool. The major challenge for this technique is the state space explosion. Since the SUT processes are running on top of the virtualization tool and the model checker must cover all the processes including virtualization tool processes, this will lead to the exponential growth of states.

Nakagawa et. al. [17] develop a model-checking framework based on this approach. Their proposed framework can execute very close to the actual model-checking execution environment. They combine the user-mode Linux and the GNU debugger (GDB) to save and

restore the entire Linux state. GBD can support several programming languages including Java. Processes are running on a virtualization tool and once non-determinism is detected within a process, the state of the OS and any possible execution paths are computed and explored by the tool.

2.2.3 Model Checker Level Centralization

The major challenges with centralization at the model checker level are managing the state space within the model checker, modeling internal communication between local threads, and possible covering of language API and classes.

A recent centralization approach has been implemented at the model checker level in the initial work [7]. They develop multi-process JVM for JPF which allows model checking of distributed Java applications. To address the problems of class confliction, static functions, and static fields, the new multi-process in JPF modifies the class loaders in JPF. The processes are mapped as a group of threads. During the initialization, each new thread is created by the SUT automatically. To capture scheduling points inside JPF, new communication models have been developed based on network API calls. This technique has been implemented into a JPF extension called jpf-nas.

Shafiei and Mehiltz [7] propose a centralization approach for model checking multiple communicating processes. This approach handles the challenge of separating data between processes by using a new class-loading model. The approach is implemented within the Java PathFinder (JPF) model checker, which allows for modeling inter-process communication (IPC) and exploring potential exceptional control flows caused by network failures. The state space of distributed applications is reduced using a partial order reduction (POR) algorithm, which is proven to preserve deadlocks. The approach also includes an automatic way to capture interactions between the system being verified and external resources.

3. Methodology

The centralization approach plays a crucial role in the model checker level and serves as the fundamental framework for JPF in model-checking distributed systems. However, for effective model-checking of these systems, it is necessary to have models that capture the communication between processes. Previous research [7] has introduced a technique to capture communication data at a byte-by-byte level by utilizing two buffers for each connection. These buffers store data that has been sent but not yet received by the other endpoint, employing a cyclic queue known as 'gov.nasa.jpf.util.ArrayByteBuffer' to store raw bytes. However, the write and read operations, responsible for adding and removing data from the queue, impose significant resource demands due to the need for backtracking through explored states during the model-checking process. As a result, this leads to a rapid expansion of the state space and computational limitations within existing IPC models.

Our work proposes a revised approach to the existing IPC models by incorporating the cache approach. Instead of handling communication data at a byte-by-byte level, the proposed method focuses on multi-byte input and output. The communication data is stored in a request and response tree, acting as a cache. During the backtracking process, the request and response pointers navigate through the data, allowing for more efficient exploration and analysis.

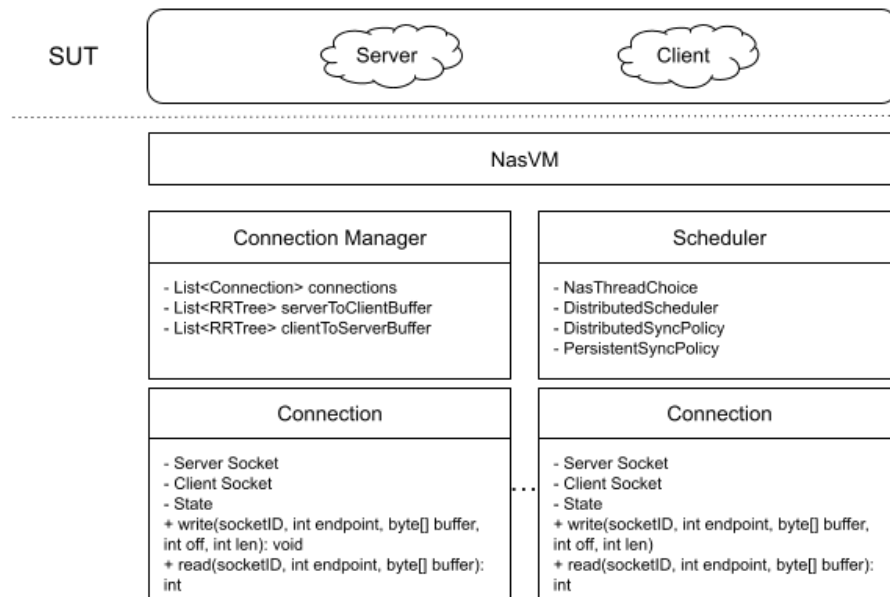


Figure 3: Overall Architecture of the proposed remodeling of IPC

Figure 3 shows the overall architecture of the proposed remodeling of IPC. The revised IPC model is an extension of JPF that can model-check distributed systems. It is a combination of the cache and centralization features. It uses the multi-process VM from the JPF core, the connection manager, the connection, and distributed scheduler. However, from the preliminary experiments, the read and write operations of the communication data byte-by-byte are not suitable for model-checking distributed systems since the states of the programs can grow exponentially and it is a resource-intensive operation during the backtracking process. The cache request and response tree are the most appropriate solution, as it only navigates request and response pointers during the backtracking process.

The following describes the proposed redesign of IPC which includes important components such as the NasVM, the connection manager, the connection, and the distributed scheduler.

- 1) **NasVM:** This class serves as the primary starting point. It begins by extending the multi-process VM from the JPF core system and then launches the connection manager component. Next, it starts the server process first and then the client process, and finally sets the global scheduling point for both processes.
- 2) **Connection Manager:** This component is created only once by the NasVM during the entire lifespan of the model checker. It is responsible for managing the connections between processes. One connection is established through two active sockets, for example, between the server and the client. The server typically starts by specifying an address and port number and then waits for the client to connect. If the client connects to the server successfully, both processes can then send and receive data from each other. When multiple threads or processes create more connections, the connection manager must manage multiple connections. The two most important variables are the buffers (communication data). In the previous centralization approach, the buffers are part of the connection object. In this work, the two buffer variables are created only once during the entire lifespan of the model checker because the data acts as a cache in the connection object. During the

backtracking process, instead of creating and deleting data, the request and response pointers are used to navigate the request and response tree. The connection manager creates two buffers, which are represented as a list type of the request and response tree (RRTree) shown in Figure 3. One buffer stores the data sent from the server to the client, and the other buffer stores the data sent from the client to the server. To keep track of the states of the connections and buffers, the connection manager has to implement a listener from the JPF core system. This is to listen to the JPF core system when it goes through bytecode executions. To achieve this, the connection manager has to implement a state extension client listener from the JPF core, which consists of three main functions: `getStateExtension()`, `restore()`, and `registerListener()`. The `getStateExtension()` function listens to the JPF core when it finds a new state, `restore()` is used when the JPF restores a state, and `registerListener()` registers the connection manager component to the state extension client.

- 3) **Connection:** The server socket object consists of a passive socket, in which the server process creates the server object and waits for the client to connect, after the successful connection, the server creates an active socket. The same thing happens to the client process, after the successful connection, the client creates an active socket. The connection manager considers this as one connection; hence, multiple server/client pairs will create multiple connections. The proposed design includes a connection object that includes a server socket, a client socket, the connection state, and read and write operations. However, the connection object cannot store the two buffers of the communication data due to the significant computation and state space explosion that occurs during the backtracking process when using the data structure. Therefore, in this design, the two buffers are left out. Instead, the read and write operations will generate two RRTrees for storing and receiving data, and utilize the request and response pointers to navigate the data. This approach helps to avoid the computational and state space issues associated with including the buffers in the connection object.
- 4) **Distributed scheduler:** JPF has a scheduler that is designed to observe the various ways in which threads are scheduled within a process. During the search for the state space of the SUT, JPF employs this scheduler to analyze and explore different sequences of concurrent transitions. These transitions could lead to varying behaviors of the process if executed in different orders. The JPF scheduler operates at a local level, which means it only deals with operations that are specific to a single process and do not involve communication between multiple processes. This approach is commonly referred to as the JPF local scheduler. For processes to communicate with one another, they must have access to a communication channel. However, the behavior of a distributed system can vary depending on the order in which processes access these channels. To account for the different orderings of concurrent transitions that involve interprocess communications, a mechanism is required. The IPC model tackles this issue by incorporating a global scheduler, which allows for the modeling of different possible orderings of communication events between processes.

3.1 The Caching Technique

Instead of storing the communication data in the “ArrayByteQueue”, the proposed method uses the request and response tree. The communication data is not deleted during the backtracking process, and it uses the request and response pointer to navigate through the data. Initially, the request and response tree (RR tree) consists solely of the root node, and when the request and response pointers are first established, they also point to this root node.

To illustrate the set of operations mentioned above, consider the following example. If the client sends “0”, the server replies “a”. After that the client sends “1”, and the server replies “b”. The first step is to create a new RR tree by invoking the main constructor and providing the socketID as a parameter. Since this is the first connection, the socketID equals zero. Then, to distinguish between a client or server connection, it is necessary to call setConnectionID(int endpoint). Assuming the client end is 567 and the server end is 763. In the beginning, the tree only contains one node, which is the root node, and to ensure that the request and response pointers are set to the root node, the addPointers() function must be called. If the target application attempts to transmit the character “0”, the client operates the addSendEvent(“0”). This operation is also addReceiveEvent(“0”) to the server tree. After that, the server operates the addSendEvent(“a”). This operation will add the “a” as the request to the server tree, and the addReceiveEvent(“a”) method adds the response “a” to the client tree as shown in Figure 4.

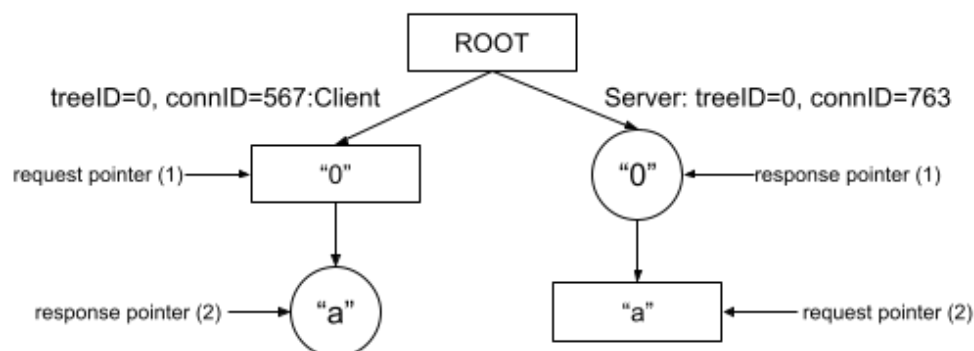


Figure 4: RR Tree Illustration After The Client Sends “0” And The Server Replies “a”

If another message in the target program tries to send another message “1” to the server, the addSendEvent(“1”) method is executed, and the addReceiveEvent(“1”) method adds the response message to the server tree. Once the transmission is complete, the RR tree for both the client and the server tree is displayed in Figure 5.

The RR tree comprises a tree ID, connection ID, a collection of nodes, and two HashMaps for tracking send and receive pointers, which are represented by a HashMap<Integer, Integer>. To ensure that each connection has a unique tree, the tree ID is assigned by the socket ID. Furthermore, the connection ID is used to differentiate between a client and server connection. To ensure that the tree is not empty, a root node is created when the RR tree is first established. After that, it is essential to specify the connection ID, which is necessary for determining whether the tree is a client or a server tree. Finally, the send and receive pointers are generated and pointed to the root node.

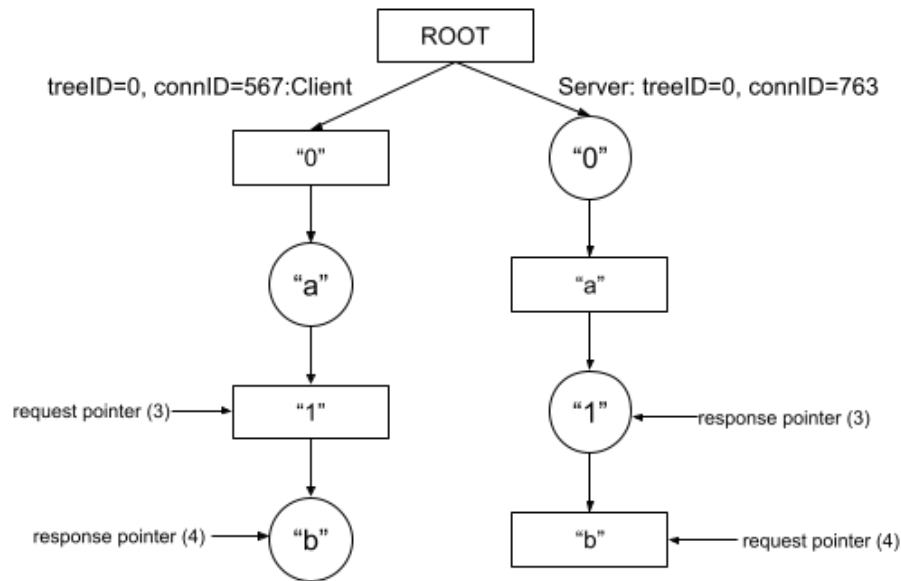


Figure 5: RR Tree Illustration After the Client Sends “1” And The Server Replies “b”

4. Experiments and Results

This section presents the results of experimenting with the IPC remodeling approach on various Java networked programs. The results are compared with a previous study [7] which utilized centralization at the model checker level. The same Java networked programs used in the study by Artho et al. are used as the test subjects. Table 1 lists these programs and includes the name of the program, its size in terms of lines of code, and its architecture. The simplest program is Echo, which has single-threaded server and client processes, while the most complex program is Alphabet, which has multi-threaded server and client processes.

Table 1 Java Networked Applications Used in The Experiments.

Application	Size (loc)	Architecture
Echo	63	Client/server
Daytime	56	Client/server
Chat	135	Client/server
Alphabet	104	Client/server

When it comes to verifying a complex application, it often involves multiple threads. The main focus is on two types of multi-threaded programs: those in which the primary thread generates all the worker threads and those in which the primary thread generates only the initial worker thread and then allows it to create additional worker threads. To manage both types of multi-threaded programs, the request and response tree can be utilized.

For each experiment, we compare the total number of states explored by the model checker, representing the state space size, the total number of bytecode instructions executed by the

JPF's JVM throughout the entire model-checking process, the maximum depth of the JPF search tree in terms of the number of transitions explored by the model checker, the elapsed time which represents the total time spent, and maximum memory which shows the maximum Java heap size.

All the experiments are implemented using the computer with the configuration described as follows:

- Operating System: Windows 10 Pro 64-bit (10.0, Build 19044)
- Processor: Intel(R) Core(TM) i7-10870H CPU @ 2.20GHz (16 CPUs), ~2.2GHz
- Memory: 16GB
- Java Development Kit: 1.8.0_333
- Java PathFinder Core System: 8.0
- Eclipse IDE: 4.14.0

4.1 Experimental Settings

To differentiate between the proposed remodeling of IPC and the existing IPC, we use JPF-Nas-Hybrid (JNH) and JPF-Nas which represents the remodification and the original IPC, respectively. Both model checkers are an extension of JPF and are configured to run in multi-process mode, employing a multiprocess virtual machine and a distributed scheduler. The distributed Java applications' original code is used as input to the model checker. The applications under the centralization model checker require the server process to begin first; otherwise, an "IOException" arises. To guarantee that executions start with a specific process, the "vm.nas.initiating_target" property is set to the server process. Also, the "vm.process_finalizers" property is set to true to allow for finalizers, which are necessary to capture certainly distributed application executions and clean up the communication channels. These two properties are used throughout the experiment.

To collect data on the decisions made during the model-checking process, we utilize a listener named "gov.nasa.jpf.listener.StateSpaceAnalyzer". This data is used to identify the primary factors that impact the size of the state space in both approaches. Additionally, we compare the scalability of the approach to the other approach based on variables such as the number of processes and the number of messages sent between processes. Finally, we provide details on the memory usage of each approach by reporting the maximum heap memory, in megabytes, that was consumed during the entire model-checking process.

Finally, for all the experiments, we only consider 20 minutes for each run, if any of the runs exceed more than 20 minutes, we report it as a state space explosion. Additionally, the existing approach can easily hit the CPU performance to 100 percent. In this case, we also report that the run reaches a state space explosion, and we will terminate the process execution.

4.2 Echo Application

The first experiment uses an application called Echo which includes a single server process and a single client process. Both processes are designed to operate using a single thread and rely on TCP sockets for communication. Once the two processes are connected, the client transmits a series of messages to the server, and the server responds by sending the same messages back to the client in the same sequence.

In this experiment, we model checks the Echo application by changing the number of threads on both the server and client. In each run, the variables “numClients” on the server code and “numThreads” on the client code are increased one by one until we reach the state space explosion.

Table 2 shows the experimental results for model checking Echo application. It shows that this approach is more effective as it leads to smaller state spaces, fewer bytecode instructions, and less complex search trees in all scenarios. The difference in state space size becomes more noticeable when there are more network interactions. This is largely due to the way we model communication channels - o approach places communication buffers as the request and the response tree and keeps them in the connection manager and it does not involve read and write operations, unlike the array byte queue from the existing method.

Table 2: Execution Results Obtained from Model Checking *Echo* Application.

Thread		JPF-NAS-HYBRID			JPF-NAS		
Client	Server	Time (second)	States	Max Memory (MB)	Time (second)	States	Max Memory (MB)
1	1	0	60	243	0	66	243
2	2	0	741	243	0	1397	243
3	3	2	8767	432	8	27288	688
4	4	28	98544	432	158	483298	2416
5	5	411	1075796	782	N/A	N/A	N/A

4.3 Daytime Application

In our second experiment, we utilized the Daytime application, which consists of a server and one or more clients that communicate with each other. Both the server and client processes are single-threaded and utilize TCP sockets for communication. When a connection is established between the server and a client, the server generates a "java.util.Date" object that captures the current time and sends its string representation to the client at that time. While a client is being served, any remaining clients are blocked until a connection with the server can be established.

In this experiment, we used our proposed model checker and JPF-Nas to perform a model check of the Daytime application. The complete source code for both the server and client can be located in Appendix A. The number of threads is incremented one by one in this scenario. If the execution time surpasses 20 minutes, we will terminate the execution.

Table 3: Execution Results Obtained from Model Checking *Daytime* Application.

Thread		JPF-NAS-HYBRID			JPF-NAS		
Client	Server	Time (second)	States	Max Memory (MB)	Time (second)	States	Max Memory (MB)
1	1	0	62	243	0	100	243
2	2	7	9145	688	23	92598	991

The results of testing the Daytime application using JNH and JPF-Nas model checkers are displayed in Table 3. The table has three columns: the first one indicates the number of threads

for the client and server, while the second and third columns present the time in seconds, the number of states, and the maximum memory used by the JNH and JPF-Nas model checkers, respectively. In every case, our proposed method results in a smaller state space, fewer bytecode instructions, and a shallower search graph. However, we can model check Daytime application up to 2 threads only due to the complexities of the application. If we set the number of client threads and the server threads to 3, the execution time will be longer than 20 minutes.

4.4 Chat Application

Our next experiment involves the Chat application, which consists of a server that communicates with one or more clients using TCP sockets. The server is designed to handle multiple threads, while the clients only use a single thread. Client interactions are managed through a worker thread dedicated to handling each interaction. When a client connects to the server, the main server thread creates a worker thread to manage its communication. This allows multiple clients to be served simultaneously. Whenever a worker thread receives a message from a client, it sends it to all the other clients in the system using a shared array that stores references to all the worker threads. The worker thread then uses these references to obtain the sockets connecting the chat server to each client.

One of the input variables for Chat is the size of the array that stores references to the worker threads. The maximum number of workers that can serve clients at the same time is equal to the size of the array. To test Chat, we conducted experiments with varying numbers of client and server threads. Table 4 displays the execution times in seconds, the number of states, and the memory consumption obtained from applying both approaches.

Table 4: Execution Results Obtained from Model Checking Chat Application.

Thread		JPF-NAS-HYBRID			JPF-NAS		
Client	Server	Time (second)	States	Max Memory (MB)	Time (second)	States	Max Memory (MB)
1	1	0	497	243	0	251	243
2	2	61	330098	687	35	200148	1095

In this experiment, we discover that our approach leads to a larger state space and longer time. However, our proposed method still consumes less memory. If we look at the source code shown in Appendix A, each client sends the letter "H" to the server, and the server forwards the letter to all clients. Since the existing centralization process data byte by byte in the array byte queue, the results lead to a smaller state space and lesser computation time compared to ours.

4.5 Alphabet Application

We conducted our last experiment using the Alphabet application, which involves a single server connected to one or more clients. Both the server and clients have multiple threads, and they use TCP sockets to communicate. Each client is assigned a string of digits, and their task is to get the server to convert the string into a string of letters. For example, the string "123" would become "ABC". To achieve this, each client creates multiple pairs of producer and consumer threads, with each pair responsible for transforming a portion of the client's string. The client determines the number of pairs and the size of the substrings to be transformed. Each producer and consumer pair connect independently to the server, and the server creates a worker thread to handle each connection. For each digit in the substring, the producer thread sends a

byte representing the digit to the server, which then converts it to a letter and sends it to the consumer thread. The client and server can handle multiple connections simultaneously, and multiple communication channels can exist between them.

The results of experimenting with the Alphabet application using JNH and JPF-Nas model checkers are displayed in Table 5.5. The table has three columns: the first one indicates the number of threads for the client and server, while the second and third columns present the time in seconds, the number of states, and the maximum memory used by the JNH and JPF-Nas model checkers, respectively. In every case, the proposed method results in a smaller state space, fewer bytecode instructions, and a shallower search graph.

Table 5: Execution Results Obtained from Model Checking *Alphabet* Application.

Thread		JPF-NAS-HYBRID			JPF-NAS		
Client	Server	Time (second)	States	Max Memory (MB)	Time (second)	States	Max Memory (MB)
1	1	0	549	243	0	100	243
2	2	61	227387	687	82	301409	1691

5. Conclusions and Future Work

Addressing the non-deterministic execution of threads, processes, and communication channels presents substantial hurdles in the development of dependable distributed systems. Our proposed approach introduces the request and response tree as a reliable storage mechanism for communication data. Additionally, our work processes data in multi-byte chunks, enabling efficient transmission between the client and server. We have also enhanced the RR tree to support multiple connections, making our method scalable as the number of RR trees aligns with the number of sockets created by the server/client, with automatic incrementation. In our approach, the nodes in the tree represent the sent or received data, and during the backtracking process, we utilize request and response pointers to traverse the tree, replacing the need for traditional write and read operations. The results show the effectiveness of our work in providing a more accurate IPC model for capturing interactions between processes. Notably, our approach successfully detected a bug that went unnoticed when employing recent centralization techniques at the model checker level. Moreover, the comprehensive evaluation demonstrated that our implementation choices significantly improved the overall performance.

Future work in this area holds promising avenues for further enhancement. One potential direction is to explore the integration of advanced optimization techniques to mitigate the state space explosion and computational limitations associated with backtracking in existing IPC models. By addressing this avenue, we can pave the way for even more robust and efficient model-checking techniques for distributed systems, enabling the development of highly reliable and secure software systems.

Acknowledgment

The authors would like to thank the Prototype Development Research Grant Scheme (PRGS) (203.PKOMP.6740069) by the Ministry of Higher Education (MOHE) for funding the research project.

References

- [1] R. Vitillo, *Understanding Distributed Systems: What every developer should know about large distributed applications*. Roberto Vitillo, 2022.
- [2] C. Chan and B. Cooper, “Debugging Incidents in Google’s Distributed Systems,” *Queue*, vol. 18, no. 2, pp. 47–66, Apr. 2020, doi: 10.1145/3400899.3404974.
- [3] A. Muscholl, “Automated Synthesis: a Distributed Viewpoint,” in *37th IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS 2017)*, 2018.
- [4] E. M. Clarke, T. A. Henzinger, and H. Veith, “Introduction to model checking,” in *Handbook of Model Checking*, Springer, 2018, pp. 1–26.
- [5] C. Baier and J.-P. Katoen, *Principles of model checking*. MIT press, 2008.
- [6] D. Beyer and A. Podelski, “Software Model Checking: 20 Years and Beyond,” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 13660 LNCS, pp. 554–582, 2022, doi: 10.1007/978-3-031-22337-2_27/TABLES/3.
- [7] N. Shafiei and P. Mehrlitz, “Extending JPF to verify distributed systems,” *ACM SIGSOFT Software Engineering Notes*, vol. 39, no. 1, pp. 1–5, Feb. 2014, doi: 10.1145/2557833.2560577.
- [8] E. Sherman et al., “The Java Pathfinder Workshop 2022,” *ACM SIGSOFT Software Engineering Notes*, vol. 48, no. 1, pp. 19–21, Jan. 2023, doi: 10.1145/3573074.3573080.
- [9] C. Artho and W. Visser, “Java Pathfinder at SV-COMP 2019 (Competition Contribution),” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11429 LNCS, pp. 224–228, 2019, doi: 10.1007/978-3-030-17502-3_18/FIGURES/1.
- [10] C. Artho, W. Leungwattanakit, M. Hagiya, and Y. Tanabe, “Efficient model checking of networked applications,” in *Lecture Notes in Business Information Processing*, Springer Verlag, 2008, pp. 22–40. doi: 10.1007/978-3-540-69824-1_3.
- [11] C. Artho, W. Leungwattanakit, M. Hagiya, Y. Tanabe, and M. Yamamoto, “Cache-based model checking of networked applications: From linear to branching time,” in *ASE2009 - 24th IEEE/ACM International Conference on Automated Software Engineering*, 2009, pp. 447–458. doi: 10.1109/ASE.2009.43.
- [12] W. Leungwattanakit, C. Artho, M. Hagiya, Y. Tanabe, and M. Yamamoto, “Model checking distributed systems by combining caching and process checkpointing,” in *2011 26th IEEE/ACM International Conference on Automated Software Engineering (ASE 2011)*, 2011, pp. 103–112.
- [13] S. D. Stoller and Y. A. Liu, “Transformations for model checking distributed Java programs,” in *International SPIN Workshop on Model Checking of Software*, 2001, pp. 192–199.
- [14] C. Artho and P. L. Garoche, “Accurate centralization for applying model checking on networked applications,” *Proceedings - 21st IEEE/ACM International Conference on Automated Software Engineering, ASE 2006*, pp. 177–186, 2006, doi: 10.1109/ASE.2006.10.
- [15] L. Ma, C. Artho, and H. Sato, “Analyzing distributed Java applications by automatic centralization,” in *2013 IEEE 37th Annual Computer Software and Applications Conference Workshops*, 2013, pp. 691–696.

- [16] E. Barlas and T. Bultan, “NetStub: A framework for verification of distributed Java applications,” *ASE '07 - 2007 ACM/IEEE International Conference on Automated Software Engineering*, pp. 24–33, 2007, doi: 10.1145/1321631.1321638.
- [17] Y. Nakagawa, R. Potter, M. Yamamoto, M. Hagiya, and K. Kato, “Model checking of multi-process applications using SBUML and GDB,” in *Proc. Workshop on Dependable Software: Tools and Methods*, 2005, pp. 215–220.
- [18] W. Leungwattanakit, C. Artho, M. Hagiya, Y. Tanabe, M. Yamamoto, and K. Takahashi, “Modular software model checking for distributed systems,” *IEEE Transactions on Software Engineering*, vol. 40, no. 5, pp. 483–501, 2014, doi: 10.1109/TSE.2013.49.

HEPATITIS C INFECTION IN HEMODIALYSIS PATIENTS, MAKKAH, SAUDI ARABIA

Ghanim Hamid Al-Khattabi¹

¹Preventive Medicine and Public Health Executive Administration, Ministry of Health, Makkah, Saudi Arabia.

Correspondence: Ghanim Hamid Al-Khattabi

Consultant of Preventive Medicine and Public Health, Epidemiologist; Preventive Medicine and Public Health Executive Administration, Ministry of Health, Makkah, Saudi Arabia.

Tel: +966570570057

Email: dralkhattabi@gmail.com

Abstract: Hepatitis C virus (HCV) infection is the most common cause of chronic liver disease in the world and evaluation of the epidemiology of HCV infection was made possible by the development of a serological assay to detect antibodies to epitopes of HCV. Hemodialysis (HD) is considered to be one of the main risk factors of HCV transmission. The prevalence of antibodies to HCV (anti-HCV) in patients undergoing maintenance dialysis therapy (MDT) was found to be one of the highest among different risk groups and there is wide variation in the prevalence of HCV infection among different dialysis units and countries. Therefore, the prevalence of anti-HCV was studied in patients undergoing MDT. The current study is a cross section one, where a sample of 361 HD patients were selected randomly from HD centers in three governmental hospitals in Makkah city and they were subjected to assessment for the prevalence of anti-HCV by using questionnaires in addition to clinical measurements. The overall prevalence of anti-HCV among HD patients was 49.9% according to clinical measurements and questionnaires. The overall prevalence of anti-HCV among HD patients in Makkah city (49.9%) was comparable to that reported from other parts of the Kingdom of Saudi Arabia (KSA) and it is almost similar to the already reported positivity rate of (43.2%) from the eastern region of KSA and southern regions of KSA (45.5%). However, Strict adherence to universal precautions as recommended by the Center for Diseases Control (CDC), meticulous regular disinfection of HD machines, Strict isolation of HCV-positive patients, dedicated dialysis machines and nursing staff at new dialysis set-up could possibly be the reasons of relatively low anti-HCV positivity at the current study than that reported by Shaheen, et al , from four centers in the western region of KSA, (72.3%) and that by Huraib, et al in their multi-center study in KSA, (68%) as well as the mean national rate. On the other hand, the use of more sensitive third generation enzyme - linked immunosorbent assay (ELISA) technique, long duration on HD, multiple blood transfusions, patients with dialysis treatment in multi-centers, factors related to infrastructure, environment and operational system might be responsible for the high prevalence of anti-HCV found in this study than that reported by Saeed, et al from Riyadh and those observed in the central region of KSA. Although routes of transmission are still unclear, early detection of all infected patients is mandatory for HCV prophylaxis in HD patients. Furthermore, an intensive educational program for staff members, HD patients and proper evaluation of the HD situation are needed. Thus, observation of appropriate preventive measures by all HD-centers is paramount.

Keywords: Hepatitis C virus, Hemodialysis, Makkah, Saudi Arabia, Antibodies Against Hepatitis c Virus (Anti-HCV)

Introduction:

Hepatitis C virus (HCV) was first identified in 1989 by Choo et al. [1] HCV infection is a persistent public health concern in hemodialysis (HD) patients [2, 3] and is the most common cause of chronic liver disease in the world [4] and in HD patients.[5] HD patients are vulnerable to HCV infection because of the risk for exposure to HCV associated with the dialysis procedure [6]. In contrast with the hepatitis B virus (HBV), no vaccine is available for HCV [7]. Accurate testing for HCV is complicated by regional variation in the HCV genome and by variation in screening tests [7-10]. Patients infected with HCV often have minimal clinical evidence of disease [7, 9, 11]. HCV infection in end-stage renal disease (ESRD) patients has been associated with greater morbidity and mortality. [9, 12, 13] HCV infection has also been associated with greater morbidity and mortality in ESRD patients after they have received a renal transplant [14-16] and is particularly highly prevalent in developing countries.[17]

There is wide variation in the prevalence of HCV infection among different dialysis units and countries as shown by Dialysis Outcomes and Practice Patterns Study (DOPPS).[18]

Approximately 170 million people in the world are infected with HCV.[19] The reported prevalence of anti-HCV antibody world-wide among dialysis patients has ranged from 1.7% [20] to 55% [21], from 1.7 to 70% [22], from 2% in northwestern Europe to 76.3% in Indonesia [23, 24] and from 29.4% to 65% in Brazil.[25-28]

However, In Kingdom of Saudi Arabia (KSA), a positivity rate of HCV ranging between 14.5 and 94.7% has been reported.[29] Furthermore, studies from dialysis units in KSA have reported prevalence rates of 34.8% to 53.7%, 46.5%, 45.5% and 72.3% from the central[30, 31] , eastern[32] , southern[33] , and western[34] regions, respectively. Other study conducted by Shaheen, et al [34] reported anti-HCV prevalence rate of 62.7% from Jeddah city and in a multi-center study of 22 HD units in KSA performed by Huraib, et al [35] reported a prevalence rate ranging from 14.5 to 94.7%.[36]

Patients on long-term HD are considered to be a population at risk for HCV infection and the prevalence of anti-HCV in patients undergoing maintenance dialysis therapy (MDT) was found to be one of the highest among different risk groups.[37] Chronic HD is associated with both endemic cases and, more rarely, sporadic outbreaks of HCV infection.[38] Therefore, the prevalence of antibodies to HCV (anti-HCV) was studied in patients undergoing MDT.

HCV infection is of great clinical importance since: (a) it is commonly associated with chronic liver disease and cirrhosis. [6][39] [40] [41] (b) it can lead to the development of hepatocellular carcinoma. [42] [43] (c) there is an increased risk of developing chronic liver disease in renal allograft recipients who are anti-HCV positive . [44] [45, 46] (d) till date, no vaccine has been developed for HCV. (e) there is an increased risk for development of chronic hepatitis in anti-HCV positive renal transplant recipients [47]. Although HCV-associated liver disease typically takes decades to become clinically manifest, a period of time much longer than the lifespan of most dialysis patients with a 5-yr survival of 60 to 70% , the liver disease–related complications seem the unlikely link to the high death risk.[48] However, the magnitude of HCV transmission within HD units is still unclear and therefore general recommendations for prevention have not been developed.[3, 49, 50] The Centre for Diseases Control (CDC) has made no recommendations for controlling HCV in HD units.[3] However, the natural course

of HCV in HD patients is not well understood. It seems to differ from that in other HCV patients.[51] Liver function tests are close to or near normal in many cases,[52, 53] but the mortality of HCV infected HD patients seems to be enhanced compared with HCV negative HD patients in preliminary studies.[54] Thus patients with HCV on chronic HD are at increased risk of death, which suggests that the focus should be directed more to identification and prevention of HCV infection in HD patients.

Despite the importance of this topic, there is remarkable paucity in the studies which identify prevalence of anti-HCV in HD patients in KSA. This paucity was detected after extensive search using the key words (Hepatitis C virus, Hemodialysis, Saudi Arabia) in both Pub Med and Google Scholar.

The aim of this study was to identify the prevalence of anti-HCV among HD patients at governmental kidney centers in Makkah city in KSA for the purpose of adopting practical measures to stop or minimize the spread of HCV infection among dialysis patients.

Materials and Methods:

The current cross-sectional study was carried out in three HD centers at three major governmental hospitals in Makkah city which is located in the western region of KSA. The HD machines are always busy through the year with four shifts daily. Sometimes a fifth shift has to be arranged to overcome the heavy workload with large influx of patients during the Holy months of Ramadan and Hajj due to a large number of visitors from outside the city of Makkah and there is an arrangement for visitors who have ESRD and need dialysis to do it in governmental hospitals. Small HD units are also available in other governmental and private hospitals, which accommodate for only a small percentage of patients.

The study population represents patients registered in the three HD centers present in three major governmental hospitals in Makkah city. Their total number was estimated to be around 2000 patients at the time of the study; they included all patients with ESRD requiring HD and they are all registered in the HD units of kidney centers. Patients at the HD centers in the three hospitals undergo HD for an average of 3 times a week, with a small percentage undergoing HD only twice a week. The HD sessions usually take place in four shifts, from 7:30 A.M to 10:30 A.M, from 11:00 A.M to 2:00 P.M, from 2:30 P.M to 5:30 P.M and from 6:00 P.M to 9:00 P.M.

The sample size needed for estimating prevalence of anti-HCV was calculated by using Epi-Info program version 6.04; the required sample size was 385 patients. The responded who completed the participation accounted for 361; making a response rate of 93.8%. Inclusion criteria were any patient, conscious, understands, able to give an informed consent (if the patient was less than 18 years, the consent was taken from his parents) and regular on HD. Stratified sampling was conducted to ensure representativeness of male and female patients. Stratification was based on the available list of patients at each HD center in each hospital. This list cover patients from all wards, including male and female wards, the isolation section, both hepatitis C positive and negative patients and at different times of the day.

As the number of patients who were available at the time of the study accounted for 770 patients, an estimated sample size was 385 which represented one half of the patients; therefore, the estimation designated sample in each place was half of available listed patients

as described in table one.

The allocation of patients from each list was done by systematic sampling selecting every second patient in each list. There were separate lists for males and females which facilitated allocation of the sample by gender.

Prevalence of anti-HCV in patients with ESRD was measured by a variety of methods. However, questionnaires and clinical measurements were used to measure prevalence of anti-HCV in HD patients.

A set of a structured questionnaire was developed by the researcher to ascertain information on patients' demographic characteristics and HCV status of patients and was translated into Arabic then it was back translated to ensure lexical equivalence. Additionally, it was subjected to validity testing after being translated into Arabic language and reviewed by consultant of family medicine, consultant of community medicine and nephrology consultant.

All patients underwent dialysis three times weekly. All units follow the same machine chemical disinfection protocols with no dialyzer reuse. The anti-HCV (Hepatitis C Virus Antibodies) assay was performed for every patient included in the study by using a third-generation enzyme - linked immunosorbent assay (ELISA) and for newly positive-case polymerase chain reaction (PCR) was performed for confirmation.

Dry weight (weight at the end of dialysis treatment) which is taken as the lowest tolerable weight at the end of dialysis treatment without the development of symptoms or hypotension, and the inter dialytic weight gain (IDWG) is calculated as the difference between the patient's weight obtained at the onset of a dialysis treatment and the weight obtained at the end of the previous dialysis. The clinical measurements included biological measurements which included in addition to interdialytic weight gain (IDWG), biochemical markers which included pre-HD serum potassium or phosphorous levels and anti-HCV assay.

Upon arrival to the dialysis centers, the physicians in charge were contacted and the researcher used to present the study design and explain the purpose of the research to the HD staff. Information regarding the HD center was obtained from the chief HD nurse. According to selection and inclusion criteria, designated patients were invited to participate in the study after explaining to them the purpose of the study.

Medical files of the patient were examined to check the weight of patients pre hemodialysis, weight of patients post hemodialysis (dry weight), number of co-morbid diseases and the presence of chronic diseases (such as DM, HTN), psychiatric diseases, hospitalization history, kidney transplant history, causes of kidney failure, hepatitis profile, potassium and phosphorus level, for how long is he or she on dialysis, number of daily tablet, others... This information were available for every patient as they are doing a monthly blood test examination pre and post hemodialysis session to evaluate the level of potassium, phosphorus and others chemical indicators.

Weight for each patient was measured before and after each HD sessions by well-trained nurses. The patient's weight at the beginning of dialysis session was subtracted from the weight at the end of previous dialysis session (dry weight) to calculate the interdialytic weight gain

(IDWG); then this IDWG is divided by weight at the end of previous dialysis session (dry weight) to get IDWG percentage. A standard electronic weighing chair was used to obtain the weight. The scale was placed on a hard floor surface. Participants were asked to remove their heavy outer garments; female patients were weighed with Abaya (ladies body cover), and Abaya was weighed and its weight was subtracted from the total. Weight was measured in all participants and taken to the nearest 0.1 kg using weighing scale. The scale was calibrated at the beginning and end of each examining day. The scale was checked using the standardized weights and calibration was corrected if the error was greater than 0.1 kg.

The pilot study was carried out in Jeddah kidney center at King Fahd general hospital (referral center for HD patients in Western province) on about 10% of our sample size who were not included in the main study.

Statistical Package for Social Science (SPSS) program, version 20 was used for statistical analysis of data. The level of statistical significance was set at $P < 0.05$ throughout the study.

Data was obtained after applying for ethical approval from the Ministry of Health & General Directorate of Health Affairs in Makkah city.

Table 1: Description of Samples and Responses from Each Hospital According to Gender:

	Al-Noor Specialist Hospital(ANSH)		King Abdul-Aziz Hospital(KAH)		King Faisal Hospital(KFH)	
Available patients	518		162		90	
Gender	Male	Female	Male	Female	Male	Female
Available patients according to gender	266	252	74	88	28	62
Required sample of patients	133	126	37	44	14	31
Responded patients	125	118	34	41	13	30

Results

The current study included 361 patients. Table 2 demonstrates that the majority of patients (78.9%) aged <65 years, with a mean age of 50.1 ± 15.8 years. There was almost equal distribution of males and females with an overwhelming majority of Saudis (93.9%). About two thirds of them were married (62.3%) and around one third (31.3%) were illiterate. The great majorities (88.9%) were unemployed and almost two thirds of the patients (60.1%) had monthly income between 1000 and 3000 SAR (Saudi Arabian Riyals). Table 3 shows that the most commonly identified causes of renal failure were diabetes mellitus (23%) and hypertension (21.9%). On the same line, it was found that the overwhelming majorities of the patients (93.9 %) were currently hypertensive and a considerable proportion (39.6%) were currently diabetic. One half of the patients (50.3%) were on dialysis for 60 months or more and the majority (77%) had previous history of hospitalization. Clinically, the average number of co-morbidities accounted for 3.0 diseases; the median interdialytic weight gain was 2.0 kg; the average number of daily tablet(s) taken by patients was 5.0. The mean levels of pre-hemodialysis serum potassium (K^+) and phosphorus (po_4) were $5.1(0.9)$ mmol/L and $5.3(1.8)$

mg/dl; respectively. Table 4 demonstrates that the overall prevalence of anti-HCV among HD patients was almost one half (49.9%) according to clinical measurements and questionnaires.

Table 2: Socio-Demographic Characteristics of The Study Group (N=361):

Characteristics	Frequency	%
Age		
Less than 65 years	285	78.9
65 years and older	76	21.1
Mean $\pm$ SD	50.1 $\pm$ 15.8	
Range	14-95	
Gender		
Male	172	47.6
Female	189	52.4
Marital status		
Married	225	62.3
Single	77	21.3
Others (divorced, widowed)	59	16.3
Nationality		
Saudi	339	93.9
Non-Saudi	22	6.1
Education level		
Illiterate	113	31.3
Primary school	89	24.7
Secondary school and above	159	44.0
Employment		
Employed	40	11.1
Not employed	321	88.9
Monthly income		
No income	30	8.3
Less than 1000 SAR*	16	4.4
From 1001 to 3000 SAR*	217	60.1
From 3001 to 6000 SAR*	57	15.8
From 6001 to 10000 SAR*	16	4.4
From 10001 to 15000 SAR*	14	3.9
From 15001 to 20000 SAR*	8	2.2
More than 20000 SAR*	3	0.8

*Saudi Arabian Riyals

Table 3: Clinical Characteristics Of The Study Group (N=361):

Table 1. Clinical Characteristics of the Study Group (n = 339).			
Characteristics		Frequency	%
Causes of renal failure	Hypertension	79	21.9
	Diabetes mellitus	83	23.0
	Glomerulonephritis	13	3.6
	Others	69	19.1
	Unknown	117	32.4
Medical history	Kidney transplant	23	6.4
	HTN	339	93.9

	D.M	143	39.6
	Psychiatric disease	37	10.2
Duration of dialysis (months)	12 months or less	57	15.9
	13 to 36 months	76	21.2
	37 to 60 months	45	12.6
	60 months or greater	180	50.3
	Median	60	
Hospitalization history		278	77.0
Number of comorbidities	Median	3.0	
Number of daily tablets	Median	5.0	
Interdialytic weight gain	Median	2.0 Kg	
Pre-hemodialysis serum potassium	Mean (SD)	5.1 (0.9) mmol/L	
	Range	2.7 – 10.2	
Pre-hemodialysis serum phosphorus	Mean (SD)	5.3 (1.8) mg/dl	
	Range	1.3 – 11.6	

Table 4: Prevalence of HCV Infection (N=361):

HCV indicators	HCV status	Frequency	%
Questionnaire	Positive	180	49.9%
	Negative	181	50.1%
Clinical measurement	Positive	180	49.9%
	Negative	181	50.1%

Discussion:

HD patients are at high risk for hepatitis viral infections especially for HCV infection, [55] which has been reported from different parts of the world as a common infection, have a high prevalence rate and now recognized as the principal cause of liver disease among HD patients. [56] [19]

Prevalence of HBV and HCV in HD patients were reported by several authors ranging between 0.9% to 21.6% for HBV [57, 58] and 3.4% to 65.8% for HCV. [59-61] There is a wide range in prevalence rates of HCV infection among HD patients in different regions of the world, ranging from 1% in the UK to more than 90% in Eastern Europe. [5, 62-69]

In the current study, we observed a high prevalence of HCV antibody among patients undergoing HD at different centers in Makah city. The overall anti-HCV prevalence among HD patients in Makah city (49.9%) was comparable to that reported from other parts of KSA. It is almost similar to the already reported positivity rate of (43.2%) from the eastern region of KSA [70] and southern regions of KSA (45.5%) [34, 71]. It is much lower than that reported from other centers in KSA. It is lower than that reported by Shaheen, et al [34, 36], from four centers in the Western region of KSA, (72.3%) and that by Huraib, et al [72] [29] in their multi-center study in KSA, (68%) as well as the mean national rate. [34, 73] However, this value is higher than those observed in the central region of KSA.

Strict adherence to universal precautions as recommended by the CDC and meticulous regular disinfection of HD machines, Strict isolation of HCV- positive patients, dedicated dialysis

machines and nursing staff at new dialysis set-up could possibly be the reasons of relatively low anti-HCV positivity at the current study.

The use of more sensitive third generation ELISA technique in our study could partially explain our higher prevalence rate than that reported earlier by Saeed, et al [30] from Riyadh. Factors related to infrastructure and operational system might be responsible for the very high prevalence of anti-HVC found in this study.

Dialysis treatment could be a specific independent risk factor for HCV transmission as cross infection and could be responsible for high prevalence of anti-HCV positivity in HD patients in the current study. The reported annual seroconversion rate of 7-9% among HD patients of KSA.[75]

The positive correlation of anti-HCV positivity with duration on HD coupled with 4.8% positivity among patients who had never been transfused, indicate strong possibility of nosocomial transmission of HCV, these observations have been reported from multiple HD centers. [5, 76, 77] Hence, the importance of strict adherence to universal infection control precautions should be emphasized in all the HD centers.

Blood transfusions could be responsible for the high prevalence of anti-HCV found in the current study. Blood transfusions can be a source for anti-HCV positivity according to reports [78] [79] though reports to the contrary also exist [80] . However, data on correlation between anti-HCV positivity and history of blood transfusion in dialysis patients is conflicting. Whereas, some workers reported positive correlation [34, 35, 67, 79] others found none.[81, 82] Nevertheless, HCV transmission through blood transfusion should be considered since the reported prevalence of anti-HCV antibody in healthy blood donors varies from 0.42% in Germany [83] up to 6.6% in China [84] .The overall prevalence of anti-HCV among voluntary blood donors in KSA is about 2.2% [85] . Thus, screening of blood donors for anti-HCV is essential and those with positive results should be rejected. Studies confirmed that administration of blood products is the main risk factor for developing hepatitis C.[5, 62, 66] [68, 86, 87] But duration of HD in patients with or without blood transfusions is also an independent risk factor.[68, 81, 86, 88] Thus patient to patient transmission during HD has been suggested.[89-93] Patient to patient transmission was prospectively proved in several incidence studies in HD patients.[91-93] Repeated blood transfusions is a risk factor for HCV transmission. However, data on correlation between anti-HCV positivity and number of blood transfusions is inconclusive. In some studies, a positive correlation with blood transfusion has been reported [67, 94] while others have refuted this possibility. [81, 82] An increased prevalence of anti-HCV positivity among patients on HD for longer periods of time has been reported.[95, 96]

Cross-infection could be responsible for the high prevalence of anti-HCV in HD patients in this study. Hence, the dialysis staff should observe all necessary precautions to prevent this. The allocation of separate HD machines for the anti-HCV positive patients could pose a real problem because of the large number of these patients. Since machine to patient transmission is not very well documented up to now, machine isolation may not be justified at the moment.

While the exact mode of HCV transmission through dialysis remains to be elucidated, it is reasonable to designate separate dialysis machines for anti-HCV positive patients. It would be

of interest to follow-up the seronegative patients to see the impact of their isolation on seroconversion rate.

Nosocomial transmission of HCV and HBV is an important contributing factor to the spread of these viruses.[97, 98] Despite various infection control procedures in HD patients prevalence and incidence rates of HBV and HCV are still significant. [99, 100] Studies have shown that strict aseptic measures can virtually eliminate HCV contamination, even in units with a high prevalence of HCV infection.[101] It is known that different methods of control, cleaning and disinfection of the HD membranes, machines, instruments and environmental surfaces may interfere with determined prevalence.[97].

Although the route of HCV transmission in HD patients is not yet fully elucidated, previous studies suggest that dialytic age and number of blood transfusions are closely associated with HCV seropositivity.[102] patients with dialysis treatment in multi-centers are also more prone to acquiring HCV infection[103]. Contaminated HD-machines, dialyzers, bloodline surfaces and hands of the caring staff and possibly, sharing of multi-dose heparin vials, are the possible factors that can contribute to the transmission of HCV among the multi-center visitors.

Non-uniform adherence to the universal precautions and disinfection procedures in different centers as well as practice of dialyzing both HCV positive and negative patients in the same room, with or without separate machines, might be the underlying reasons to make these patients more prone to acquire HCV. However, dialysis of HCV positive patients in a separate designated room is not recommended by the CDC.[104] Multivariate analysis has revealed that seropositivity of HCV increases with the increased dialytic age of the patient and the number of blood transfusions. However, reports to the contrary have also been published regarding the association of blood transfusion and HCV transmission.[105, 106] Increased nosocomial transmission due either to prolonged immunocompromized state or breaches in the universal infection control measures during dialytic age, may be the factors responsible for this. Recently, nosocomial transmission of HCV among HD patients has been documented by the molecular analysis,[107, 108] which raises the question of how HCV- positive patients should be handled in dialysis units. Although, it has been showed that 50-70% of HCV-positive patients progress, over a period of 10-40 years, to chronic hepatitis with increased risk of cirrhosis, liver failure and liver cancer,[43] the longterm outcome of untreated patients chronic HCV infection having normal or elevated liver enzymes is still unknown.

Conclusion:

In conclusion, this cross-sectional study was designed to identify the prevalence of anti-HCV in HD patients in Makkah city since they are at high risk for HCV infection. The overall prevalence of anti-HCV among HD patients in Makkah city (49.9%) was comparable to that reported from other parts of KSA and it is almost similar to the already reported positivity rate of (43.2%) from the eastern region of KSA [70] and southern regions of KSA (45.5%). [34, 71] However, Strict adherence to universal precautions as recommended by the CDC, meticulous regular disinfection of HD machines, Strict isolation of HCV-positive patients, dedicated dialysis machines and nursing staff at new dialysis set-up could possibly be the reasons of relatively low anti-HCV positivity at the current study than that reported by Shaheen, et al [34, 36] , from four centers in the Western region of KSA, (72.3%) and that by Huraib, et al [72] [29] in their multi-center study in KSA, (68%) as well as the mean national rate.[34, 73] On the other hand, the use of more sensitive third generation ELISA technique, long duration

on HD, multiple blood transfusions, patients with dialysis treatment in multi-centers, factors related to infrastructure, environment and operational system might be responsible for the high prevalence of anti-HCV found in this study than that reported earlier by Saeed, et al [30] from Riyadh and those observed in the central region of KSA.[74]

Recommendations:

This study recommends health leaders to improve the economic status of HD patients who need that support because almost two thirds of the patients (60.1%) had monthly income between 1000 and 3000 SAR. Since the great majorities (88.9%) of HD patients at the HD centers are unemployed, a subsidized support system must be initiated to support these patients; Social agencies and charitable organizations can collaborate with HD centers in this effort.

As the current study showed a high prevalence of anti-HCV among HD patients [almost one half (49.9%) of them were anti-HCV positive] in Makkah city, frequent measurement of prevalence of anti-HCV for them should be carried out in all HD centers in Makkah city and other parts of KSA.

Further qualitative and quantitative researches into effective ways are needed by using different research designs to identify the prevalence of anti-HCV among HD patients in Makkah city and other parts of KSA. The results of these researches are better to be compared with results of this study. These researches must continue in order to improve outcomes among HD patients.

Efforts should be done to identify factors associated with HCV infection in Makkah city and other parts of KSA in order to decrease them as those factors constitutes a major health problem for HD patients. However, in most European countries, consistent reinforcement of hygienic precautions and/or isolation strategies in HD units has resulted in a substantial decrease of both the incidence and prevalence of HCV infection in HD units [109]. So, we must concentrate more on those factors in addition to factors related to infrastructure, environment and operational system.

More detailed studies are required to know the exact mode of spread of HCV so that appropriate measures may be taken to prevent this serious disease from spreading. However, an intensive educational program for staff members and proper evaluation of the HD situation are needed. Thus, observation of appropriate preventive measures by all the HD-centers is paramount.

Strict adherence to the universal precautions combined with proper disinfection procedures and increasing awareness among the HD staff and patients may reduce the spread of HCV infection in the dialysis units. Furthermore, as routes of transmission are still unclear, early detection of all infected patients is mandatory for HCV prophylaxis in HD patients. This early detection of HCV infection is important in patients undergoing HD because of the high prevalence of infection among them. [110]. This early detection could result in better management of patients and a reduction in patient-to patient transfer of HCV infection in HD units.[110]

References:

1. Choo, Q.L., et al., *Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome*. Science, 1989. **244**(4902): p. 359-62.

2. Control, C.f.D. and Prevention, *Recommendations for preventing transmission of infections among chronic hemodialysis patients*. MMWR, 2001. **50**: p. 1-43.
3. De Medina, M., et al., *Prevalence of hepatitis C and G virus infection in chronic hemodialysis patients*. American journal of kidney diseases, 1998. **31**(2): p. 224-226.
4. Alter, M.J. and H.S. Margolis, *Recommendations for prevention and control of hepatitis C virus (HCV) infection and HCV-related chronic disease*. 1998.
5. Knudsen, F., et al., *Hepatitis C in dialysis patients: Relationship to blood transfusions, dialysis and liver disease*. Kidney international, 1993. **43**(6): p. 1353-1356.
6. Cuthbert, J.A., *Hepatitis C*. Am J Med Sci, 1990. **299**(5): p. 346-55.
7. Lauer, G.M. and B.D. Walker, *Hepatitis C virus infection*. New England journal of medicine, 2001. **345**(1): p. 41-52.
8. Bukh, J., et al., *High prevalence of hepatitis C virus (HCV) RNA in dialysis patients: failure of commercially available antibody tests to identify a significant number of patients with HCV infection*. Journal of Infectious Diseases, 1993. **168**(6): p. 1343-1348.
9. Fabrizi, F., F.F. Poordad, and P. Martin, *Hepatitis C infection and the patient with end-stage renal disease*. Hepatology, 2002. **36**(1): p. 3-10.
10. Bassit, L., et al., *Comparative study of two anti-HCV screening tests in a large genotyped population of Brazilian dialysis patients*. European Journal of Clinical Microbiology & Infectious Diseases, 2002. **21**(5): p. 404-406.
11. Fabrizi, F., et al., *Virological characteristics of hepatitis C virus infection in chronic hemodialysis patients: a cross-sectional study*. Clinical nephrology, 1995. **44**(1): p. 49-55.
12. Nakayama, E., et al., *Prognosis of anti-hepatitis C virus antibody-positive patients on regular hemodialysis therapy*. Journal of the American Society of Nephrology, 2000. **11**(10): p. 1896-1902.
13. Maisonneuve, P., et al., *Cancer in patients on dialysis for end-stage renal disease: an international collaborative study*. The Lancet, 1999. **354**(9173): p. 93-99.
14. Hanafusa, T., et al., *Retrospective study on the impact of hepatitis C virus infection on kidney transplant patients over 20 years*. Transplantation, 1998. **66**(4): p. 471-476.
15. Mathurin, P., et al., *Impact of hepatitis B and C virus on kidney transplantation outcome*. Hepatology, 1999. **29**(1): p. 257-263.
16. K.V. Rao, J.M., *Chronic viral hepatitis enhances the risk of infection but not acute rejection in renal transplant recipients*. Transplantation, 1996. **62**: p. 5.
17. Alter, H.J., *Transmission patterns in hepatitis C virus infection*, in *Viral Hepatitis and Liver Disease*. 1994, Springer. p. 445-449.
18. Young, E.W., et al., *The dialysis outcomes and practice patterns study (DOPPS): an international hemodialysis study*. Kidney International, 2000. **57**: p. S74-S81.
19. Fabrizi, F., et al., *Biological dynamics of viral load in hemodialysis patients with hepatitis C virus*. American journal of kidney diseases, 2000. **35**(1): p. 122-129.
20. Conlon, P., et al., *Lower prevalence of anti-hepatitis C antibody in dialysis and renal transplant patients in Ireland*. Irish journal of medical science, 1993. **162**(4): p. 145.
21. Mondelli, M., et al., *High prevalence of antibodies to hepatitis C virus in hemodialysis units using a second generation assay*. Nephron, 1992. **61**(3): p. 350-351.
22. Othman, B. and F. Monem, *Prevalence of Antibodies to Hepatitis C Virus among Hemodialysis Patients in Damascus, Syria*. Infection, 2001. **29**(5): p. 262-265.
23. Handajani, R., et al., *Differential prevalence of hepatitis C virus subtypes in healthy blood donors, patients on maintenance hemodialysis, and patients with hepatocellular carcinoma in Surabaya, Indonesia*. Journal of clinical microbiology, 1996. **34**(12): p. 2875-2880.

24. Schneeberger, P.M., et al., *Hepatitis C virus infections in dialysis centers in The Netherlands: a national survey by serological and molecular methods*. Journal of clinical microbiology, 1998. **36**(6): p. 1711-1715.
25. Karohl, C., et al., *Prevalência de anticorpos anti-vírus da hepatite C em pacientes em hemodiálise crônica de Porto Alegre*. J Bras Nefrol, 1995. **17**(1): p. 40-46.
26. Vanderborght, B., et al., *High prevalence of hepatitis C infection among Brazilian hemodialysis patients in Rio de Janeiro: a one-year follow-up study*. REVISTA-Instituto De Medicina Tropical De Sao Paulo, 1995. **37**: p. 75-75.
27. Naghettini, A.V., et al., *Soroprevalência do vírus da hepatite C na população em diálise de Goiânia, GO*. Revista da Sociedade Brasileira de Medicina Tropical, 1997. **30**(2): p. 113-117.
28. Góngora, D., *Marcadores sorológicos da infecção pelo vírus da hepatite C em trabalhadores e pacientes da Unidade de Diálise do Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo*. Revista da Sociedade Brasileira de Medicina Tropical, 1998. **31**: p. 585-586.
29. Huraib S, A.-R.R., Aldress A, et al, *High prevalence and risk factors for hepatitis C in Saudi Arabia*. Saudi Kidney Dis Transplant Bull, 1993. **4**.
30. Saeed, A.A., et al., *HEPATITIS-C VIRUS IN SAUDI-ARABIA-A PRELIMINARY SURVEY*. Saudi Medical Journal, 1990. **11**(4): p. 331-332.
31. Fakunle, Y.M., et al., *Prevalence of antibodies to hepatitis C virus in hemodialysis patients in Riyadh*. Annals of Saudi medicine, 1991. **11**(5): p. 504-506.
32. Fathalla, S. and A. Aljama, *Prevalence Of Hepatitis-C Viral Antibodies In Blood-Donors, Pregnant-Women And Hemodialysis-Patients In The Eastern Province Of Saudi-Arabia-A Preliminary-Study*. Saudi Medical Journal, 1993. **14**(3): p. 265-265.
33. Nasser, M.A., et al., *Seropositivity to hepatitis C virus in Saudi haemodialysis patients*. Vox sanguinis, 1992. **62**(2): p. 94-97.
34. Shaheen, F.A., et al., *Prevalence of hepatitis C antibodies among hemodialysis patients in the Western province of Saudi Arabia*. Saudi Journal of Kidney Diseases and Transplantation, 1995. **6**(2): p. 136.
35. Huraib, S., et al., *High prevalence and risk factors for hepatitis C in Saudi Arabia: a need for new strategies in dialysis practice*. Nephrol Dial Transplant, 1995. **10**(4): p. 470-4.
36. Al Shohaib, S.S., et al., *The prevalence of hepatitis C virus antibodies among hemodialysis patients in Jeddah area, Saudi Arabia*. Saudi Journal of Kidney Diseases and Transplantation, 1995. **6**(2): p. 128.
37. Muller, G.Y., et al., *Risk factors for dialysis-associated hepatitis C in Venezuela*. Kidney international, 1992. **41**(4): p. 1055-1058.
38. Marschall S. Runge, C.P., *Principles of Molecular Medicine*. 2 ed. 18-11-2007: Springer Science & Business Media. 1268.
39. Debure, A., et al., *Chronic hepatitis in kidney allograft recipients*. The Lancet, 1990. **335**(8694): p. 878-880.
40. Bruguera, M., et al., *Incidence and features of liver disease in patients on chronic hemodialysis*. Journal of clinical gastroenterology, 1990. **12**(3): p. 298-302.
41. Cuthbert, J.A., *Hepatitis A: old and new*. Clinical microbiology reviews, 2001. **14**(1): p. 38-58.
42. Kew, M.C., et al., *Hepatitis C virus antibodies in southern African blacks with hepatocellular carcinoma*. Lancet, 1990. **335**(8694): p. 873-4.
43. Kew, M., et al., *Hepatitis C virus antibodies in southern African blacks with hepatocellular carcinoma*. The Lancet, 1990. **335**(8694): p. 873-874.

44. Wright, T.L., *Hepatitis C virus infection and organ transplantation*. Prog Liver Dis, 1993. **11**: p. 215-30.
45. Wright, T., *Hepatitis C virus infection and organ transplantation*. Progress in liver diseases, 1993. **11**: p. 215-230.
46. Roth, D., et al., *Detection of hepatitis C virus infection among cadaver organ donors: evidence for low transmission of disease*. Annals of internal medicine, 1992. **117**(6): p. 470-475.
47. Chan, T.-M., et al., *A prospective study of hepatitis C virus infection among renal transplant recipients*. Gastroenterology, 1993. **104**(3): p. 862-868.
48. Kalantar-Zadeh, K., et al., *Hepatitis C Virus and Death Risk in Hemodialysis Patients*. journal of the american society of nephrology, 2007. **18**(5): p. 1584.
49. Moyer LA, A.M., *Hepatitis C virus in the hemodialysis setting: a review with recommendations for control*. Semin Dial, 1994. **7**: p. 124-127.
50. Blumberg, A., C. Zehnder, and J.J. Burckhardt, *Prevention of hepatitis C infection in haemodialysis units. A prospective study*. Nephrol Dial Transplant, 1995. **10**(2): p. 230-3.
51. Simon, N., et al., *A twelve year natural history of hepatitis C virus infection in hemodialyzed patients*. Kidney Int, 1994. **46**(2): p. 504-11.
52. Guh, J.Y., et al., *Impact of decreased serum transaminase levels on the evaluation of viral hepatitis in hemodialysis patients*. Nephron, 1995. **69**(4): p. 459-65.
53. Fabrizi F, L.G., Andrulli S, et al., *Influence of serum HCV RNA upon aminotransferase activity in chronic dialysis patients*. Nephrol Dial Transplant, 1997. **12**: p. 1295-300.
54. Stehman-Breen, C.O., et al., *Risk of death among chronic dialysis patients infected with hepatitis C virus*. Am J Kidney Dis, 1998. **32**(4): p. 629-34.
55. Hosseini-Moghaddam, S.M., et al., *Distribution of hepatitis C virus genotypes among hemodialysis patients in Tehran—a multicenter study*. Journal of medical virology, 2006. **78**(5): p. 569-573.
56. Cotler, S.J., et al., *Characteristics of hepatitis C in renal transplant candidates*. Journal of clinical gastroenterology, 2002. **35**(2): p. 191-195.
57. Schneeberger, P.M., et al., *The prevalence and incidence of hepatitis C virus infections among dialysis patients in the Netherlands: a nationwide prospective study*. The Journal of infectious diseases, 2000. **182**(5): p. 1291-1299.
58. Todorov, V., et al., *High prevalence of hepatitis c virus in one dialysis center in bulgaria*. Nephron, 1998. **79**(2): p. 222.
59. Tokars, J.I., et al. *National surveillance of dialysis-associated diseases in the United States, 2000*. in *Seminars in dialysis*. 2002. Wiley Online Library.
60. Vladutiu, D., et al., *Infections with hepatitis B and C viruses in patients on maintenance dialysis in Romania and in former communist countries: yellow spots on a blank map?* Journal of viral hepatitis, 2000. **7**(4): p. 313-319.
61. Somi, M.H., et al., *Hepatitis C virus genotypes in patients with end-stage renal disease in East Azerbaijan, Iran*. Saudi Journal of Kidney Diseases and Transplantation, 2008. **19**(3): p. 461.
62. Alter, M.J. *Epidemiology of hepatitis C in the West*. in *Seminars in liver disease*. 1995. © 1995 by Thieme Medical Publishers, Inc.
63. Chan, T.M., et al., *Prevalence of hepatitis C virus infection in hemodialysis patients: a longitudinal study comparing the results of RNA and antibody assays*. Hepatology, 1993. **17**(1): p. 5-8.

64. Courouce, A.-M., et al., *Hepatitis C virus (HCV) infection in haemodialysed patients: HCV-RNA and anti-HCV antibodies (third-generation assays)*. Nephrology Dialysis Transplantation, 1995. **10**(2): p. 234-239.
65. McIntyre, P., et al., *Hepatitis C virus infection in renal dialysis patients in Glasgow*. Nephrology Dialysis Transplantation, 1994. **9**(3): p. 291-295.
66. Medin, C., et al., *Seroconversion to hepatitis C virus in dialysis patients: a retrospective and prospective study*. Nephron, 1993. **65**(1): p. 40-45.
67. Dentico, P., et al., *Prevalence and incidence of hepatitis C virus (HCV) in hemodialysis patients: study of risk factors*. Clinical nephrology, 1992. **38**(1): p. 49-52.
68. Jadoul, M., C. Cornu, and C.v.Y. de Strihou, *Incidence and risk factors for hepatitis C seroconversion in hemodialysis: A prospective study*. Kidney international, 1993. **44**(6): p. 1322-1326.
69. Covic, A., et al., *Hepatitis virus infection in haemodialysis patients from Moldavia*. Nephrology, dialysis, transplantation: official publication of the European Dialysis and Transplant Association-European Renal Association, 1999. **14**(1): p. 40-45.
70. Al-Muhanna, F.A., *Hepatitis C virus infection among hemodialysis patients in the eastern region of Saudi Arabia*. Saudi Journal of Kidney Diseases and Transplantation, 1995. **6**(2): p. 125.
71. Saxena, A.K., et al., *Prevalence of hepatitis C antibodies among hemodialysis patients in Al-Hasa region of Saudi Arabia*. Saudi Journal of Kidney Diseases and Transplantation, 2001. **12**(4): p. 562.
72. Huraib, S., et al., *High prevalence and risk factors for hepatitis C in Saudi Arabia: a need for new strategies in dialysis practice*. Nephrol Dial Transplant, 1995. **10**(4): p. 470-474.
73. Khan, L. and S. Khan, *Prevalence of hepatitis B and C markers in patients on maintenance hemodialysis in Najran*. Saudi medical journal, 2001. **22**(7): p. 641-642.
74. Anonymous, *Dialysis in the Kingdom of Saudi Arabia (SCOT data)*. Saudi J Kidney Dis Transplant 2001. **12**(3): p. 14.
75. Souqiyyeh, M.Z., et al., *The annual incidence of seroconversion of antibodies to the hepatitis C virus in the hemodialysis population in Saudi Arabia*. Saudi Journal of Kidney Diseases and Transplantation, 1995. **6**(2): p. 167.
76. Teruel, J., et al., *Importance of nosocomial transmission of hepatitis C virus infection in dialysis units*. Clinical nephrology, 1992. **38**(4): p. 232.
77. Favero MS, A.M., Bland LA., *Dialysis associated with infections and their control*. In: *Hospital infection* 3rd Edition ed. 1992.
78. Pauri, P., et al., *Risk factors and clinical expression of HCV infection in hemodialysis patients*. Nephron, 1992. **61**(3): p. 313-314.
79. Van der Poel, C., et al., *Infectivity of blood seropositive for hepatitis C virus antibodies*. The Lancet, 1990. **335**(8689): p. 558-560.
80. Pluvio, M., et al., *Hepatitis C virus-related acute and chronic hepatitis in hemodialysis patients*. Nephron, 1992. **61**(3): p. 322-323.
81. Hardy, N., et al., *Antibody to hepatitis C virus increases with time on hemodialysis*. Clinical nephrology, 1992. **38**(1): p. 44-48.
82. Vitale, C., et al., *Epidemiology of hepatitis C virus infection in dialysis units: first-versus second-generation assays*. Nephron, 1993. **64**(2): p. 315.
83. Kuhl, P., et al., *Antibody to hepatitis C virus in German blood donors*. Lancet, 1989. **2**(8658): p. 324-324.
84. Tao, Q., et al., *Investigation of anti-HCV in 391 serum samples in China*. Chinese medical journal, 1990. **103**(8): p. 616-618.

85. Al-Mofarreh, M., et al., *Prevalence of antibodies to hepatitis C virus in blood donors in Riyadh*. Annals of Saudi medicine, 1991. **11**(5): p. 501-503.
86. Di Lallo, D., et al., *Risk factors of hepatitis C virus infection in patients on hemodialysis: a multivariate analysis based on a dialysis register in Central Italy*. European journal of epidemiology, 1999. **15**(1): p. 11-14.
87. Salama, G., et al., *Hepatitis C virus infection in French hemodialysis units: a multicenter study*. Journal of medical virology, 2000. **61**(1): p. 44-51.
88. Okuda, K., et al., *Mode of hepatitis C infection not associated with blood transfusion among chronic hemodialysis patients*. Journal of hepatology, 1995. **23**(1): p. 28-31.
89. Fabrizi, F., et al., *Acquisition of hepatitis C virus in hemodialysis patients: a prospective study by branched DNA signal amplification assay*. American journal of kidney diseases, 1998. **31**(4): p. 647-654.
90. Fabrizi, F., et al., *Detection of de novo hepatitis C virus infection by polymerase chain reaction in hemodialysis patients*. American journal of nephrology, 1999. **19**(3): p. 383-388.
91. Irish, D.N., et al., *Identification of hepatitis C virus seroconversion resulting from nosocomial transmission on a haemodialysis unit: implications for infection control and laboratory screening*. Journal of medical virology, 1999. **59**(2): p. 135-140.
92. Katsoulidou A, P.D., Kalapothaki V, et al., *Molecular epidemiology of a hepatitis C outbreak in a haemodialysis unit. Multicentre haemodialysis cohort study on viral hepatitis*. Nephrol Dial Transplant 1999. **14**(11): p. 7.
93. Mizuno, M., et al., *Genetic and serological evidence for multiple instances of unrecognized transmission of hepatitis C virus in hemodialysis units*. Journal of clinical microbiology, 1998. **36**(10): p. 2926-2931.
94. Ayoola, E., et al., *Prevalence and significance of antibodies to hepatitis C virus among Saudi haemodialysis patients*. Journal of medical virology, 1991. **35**(3): p. 155-159.
95. Giammaria, U., et al., *HCV infection in hemodialyzed patients: incidence and correlation with dialytic age*. Nephron, 1992. **61**(3): p. 335-336.
96. OGUCHI, H., et al., *Prevalence of anti-HCV in patients on long-term hemodialysis*. The Japanese Journal of Nephrology, 1990. **32**(3): p. 313-317.
97. Busek, S.U., et al., *Hepatitis C and hepatitis B virus infection in different hemodialysis units in Belo Horizonte, Minas Gerais, Brazil*. Memórias do Instituto Oswaldo Cruz, 2002. **97**(6): p. 775-778.
98. Carneiro, M.A., et al., *Hepatitis C prevalence and risk factors in hemodialysis patients in Central Brazil: a survey by polymerase chain reaction and serological methods*. Memórias do Instituto Oswaldo Cruz, 2001. **96**(6): p. 765-769.
99. Souza, K.P., et al., *Hepatitis B and C in the hemodialysis unit of Tocantins, Brazil: serological and molecular profiles*. Memórias do Instituto Oswaldo Cruz, 2003. **98**(5): p. 599-603.
100. Silva, L.K., et al., *Prevalence of hepatitis C virus (HCV) infection and HCV genotypes of hemodialysis patients in Salvador, Northeastern Brazil*. Brazilian journal of medical and biological research, 2006. **39**(5): p. 595-602.
101. Pereira, B.J., *Hepatitis C virus infection in dialysis: a continuing problem*. Artificial organs, 1999. **23**(1): p. 51-60.
102. Olmer, M., D. Bouchouareb, and C. Zandotti, *Transmission of the hepatitis C virus in an hemodialysis unit: evidence for nosocomial infection*. Clinical nephrology, 1997. **47**(4): p. 263-270.

103. Kashem, A., et al., *Hepatitis C virus among hemodialysis patients in Najran: prevalence is more among multi-center visitors*. Saudi Journal of Kidney Diseases and Transplantation, 2003. **14**(2): p. 206.
104. Saxena, A.K. and B. Panhotra, *Nosocomial transmission of hepatitis C virus: impact of strict isolation on annual seroconversion rate in a hemodialysis unit*. Saudi Journal of Kidney Diseases and Transplantation, 2002. **13**(2): p. 186.
105. Lin, H.-H., et al., *Prevalence of antibodies to hepatitis C virus in the hemodialysis unit*. American journal of nephrology, 1991. **11**(3): p. 192-194.
106. Dandrani S, H.N., Danielson S, Wilson B., *Risk of hepatitis C increases with duration of dialysis (abstract)*. J Am Soc Nephrol, 1991. **2**.
107. de Lamballerie, X., et al., *Nosocomial transmission of hepatitis C virus in haemodialysis patients*. Journal of medical virology, 1996. **49**(4): p. 296-302.
108. Grethe, S., et al., *Molecular epidemiology of an outbreak of HCV in a hemodialysis unit: Direct sequencing of HCV-HVR1 as an appropriate tool for phylogenetic analysis*. Journal of medical virology, 2000. **60**(2): p. 152-158.
109. Jadoul, M. and G. Barril, *Hepatitis C in hemodialysis: epidemiology and prevention of hepatitis C virus transmission*, in *Hepatitis C in Renal Disease, Hemodialysis and Transplantation*. 2012, Karger Publishers. p. 35-41.
110. Méndez-Sánchez, N., et al., *Prevalence of hepatitis C virus infection among hemodialysis patients at a tertiary-care hospital in Mexico City, Mexico*. Journal of clinical microbiology, 2004. **42**(9): p. 4321-4322.

DEVOPS ADOPTION GUIDELINES, CHALLENGES, AND BENEFITS: A SYSTEMATIC LITERATURE REVIEW

Usman Hamza ^{1*}
Nasuha Lee Abdullah ²
Sharifah Mashita Syed-Mohamad ³

¹ School of Computer Sciences, Universiti Sains Malaysia 11800 USM Penang, Malaysia

² School of Computer Sciences, Universiti Sains Malaysia 11800 USM Penang, Malaysia

³ Faculty of Ocean Engineering Technology and Informatics, Universiti Malaysia Terengganu, 21030 Kuala Nerus, Terengganu, Malaysia

Abstract: *The DevOps methodology in software engineering is currently trending and many software development organisations are adopting it to stay competitive. However, these organisations face difficulties in adopting DevOps due to a lack of common standard guidelines and methodology. We argued that guidelines across multiple platforms are needed to succeed in DevOps implementation. This paper aims to identify the benefits, challenges, and guidelines associated with DevOps adoption using a systematic literature review. The search spanned over ten digital libraries such as ACM Digital Library, ScienceDirect, EBSCOhost, SAGE Journals, Wiley Online Library, and so on. The review also included keyword searches in relevant studies i.e DevOps guidelines, DevOps and guidelines, DevOps benefits, DevOps and benefits, DevOps challenges, DevOps and challenges. Papers written in English and published from 2011 through July 2020 were considered. 33 studies passed the two-step quality assessment. Finally, the data were extracted and analysed based on the themes and research questions, revealing the benefits and challenges of DevOps adoption guidelines. The findings show that improved software delivery speed increases code quality and application quality with frequent releases. It also shows that there was no transparent scientific approach to determining the DevOps capabilities regarding the process, practice, tools, and implementations, cost implications, and insufficient knowledge of the automation process. Finally, the reviewed studies indicated that software quality increases whenever DevOps best practices are adopted. DevOps is about continuous software delivery. Previous studies have not considered the guidelines for scientifically adopting DevOps for organisations that want to migrate. The present review identified the DevOps adoption guideline model represented in a cyclic format that indicates an infinite process representing continuous integration, continuous delivery, and continuous deployment, with an understanding of benefits and challenges.*

Keywords: *DevOps Guidelines, Challenges, Benefits, Systematic Literature Review*

1. Introduction

DevOps originated in agile software development [1] by combining development and operations (developers and operators). DevOps is considered a phenomenon or set of principles and practices that enable collaborations and communications to produce an efficient team of developers and operators [2]. DevOps is seen as a modern software development methodology and a response to the inter-dependencies between development and operations by taking a modern approach and producing tools through collaboration between developers and operators

[3], [4]. The DevOps methodology in software engineering is evolving rapidly due to customers' high expectations and technological advancements. The current trend has motivated software development organisations and teams to adopt DevOps principles to stay competitive. However, these organisations face difficulties in adopting these principles. DevOps implementation strategies lack common standard guidelines and methodology [5], [6].

Researchers have been actively investigating DevOps' scientific and technical stand for organisations' adoption and practices and have developed many techniques and models [7], [8]. However, none of such studies focuses on specific aspects related to DevOps adoption guidelines to our best knowledge. Therefore, this reviews explores some considerations and procedures for DevOps adoption and identifies challenges and obstacles that prevent its optimal adoption. Forward/backward snowballing iterations were used to ensure that all papers matched keywords. As a result, 810 primary studies were scrutinised, of which 33 preliminary studies eventually passed the selection criterion and were extensively studied. A classification framework was developed based on the main components of DevOps methodologies.

Based on the research [9] conducted a case study using a web survey on the DevOp adoption process to characterise relevant factors of successful DevOps adoption. The reported findings [10] from the web survey analysis could serve as inputs for future framework development. However, several studies with existing methods and software development tools could benefit from improvements in the deployment [11], delivery, and maintenance phases. DevOps improves these phases by collecting principles and practices [12] and collaborating between development and operations teams. DevOps is the collaboration between development and operations teams and represents a new way of thinking in the context of the software engineering field that has recently received much attention [13]. A similar study conducted systematic mapping and selected 44 articles that discuss DevOps' definition based on DevOps practices and how DevOps is related to other existing methods. Some studies mentioned a combination of descriptions, techniques, and relations to other forms, leading other researchers to propose a definition for DevOps that includes its benefits. Despite an increasing demand for software delivery, both in industry and academia, deployment obstacles often impede the quick delivery of the new and improved features to end users[14]. Similar studies [5] have adopted an SLR approach to identify challenges to adopting DevOps in practice before disembarking on the journey effectively. There is a need to clearly understand DevOps concepts, training, tools, and benefits. The present study applied the SLR method, and the results indicate that the techniques could create shorter lead times and better communication within and between development teams[15]. There are also some challenges and drawbacks to change management, software quality assurance, and stakeholder acceptance[16], [17]. Other studies suggest considering the comprehensive reporting of context and how to implement the practice, and the adoption of a metric is very relevant for achieving an efficient security approach[18] instead of referring to the use of the procedure[17], [19]. Some studies have classified DevOps based on principles, practices, benefits, and challenges, which were not sufficiently addressed in the peer-reviewed literature[19]. This review further conducts an SLR to identify the benefits and challenges of DevOps adoption and the guidelines for comprehending the current difficulties in adoption. None of the reviewed studies indicates interest or signs of developing policies for DevOps adoption. Our study identified some common guidelines that organisations need to consider before adopting DevOps that would help them succeed in implementing DevOps.

DevOps is a collection of processes and philosophies that include the four fundamental elements of culture, collaboration, tools, and operations. Due to a lack of common standard guidelines and methodology, as well as a transparent scientific approach to determining DevOps capabilities in terms of process, practise, tools, and implementations, cost implications and a lack of knowledge of the automation process, organisations are having difficulty adopting DevOps. Using a systematic literature review, this paper aims to identify the benefits, challenges, and proposed guidelines associated with DevOps adoption. The findings show an increased software delivery speed with improved code and application quality with frequent releases when best practices were implemented. The reviewed study also found that when DevOps best practices are implemented, software quality improves. Continuous software delivery is the focus of DevOps.

2. Methodology

This study follows the guidelines for systematic reviews in software engineering research[17], [20]. The methodology has three phases: planning the review, conducting the study, and reporting the extracted results. The literature search involved an automated search of ten popular digital libraries and several iterations of the snowballing approach [21] to ensure the completeness of results.

2.1 Planning the Review

The plan is to discover the current DevOps adoption guidelines and review the existing approach. In particular, the research questions addressed by this review are:

- RQ1 What are the DevOps guidelines?
- RQ2 What are the challenges in DevOps adoption?
- RQ3 What are the benefits of DevOps adoption?

2.2 Conducting the Review

The review started with the identification of relevant primary studies. The search spanned over the 10 digital libraries most commonly used in similar studies: SCOPUS, EBSCO-Discovery Service, ACM Digital Library, ScienceDirect, Wiley Online Library, SAGE Journal, Google Scholar, SpringerLink, EBSCO Academic Search Complete, and ProQuest. The process was carried out in the five steps described in Table 1.

Table 1 Inclusion and Exclusion Criteria

Step	Inclusion Criteria	Exclusion Criteria
1	Studies presenting DevOps concepts/definitions and methodology	Studies not explicitly providing solutions to DevOps challenges and practice
2	Studies presenting DevOps strategies/practice	Studies presenting surveys
3	Studies presenting DevOps issues/goals, models, and their benefits	Editorials, abstract, thesis panels, monographs, books, book sections
4	Studies presenting DevOps principles	Articles that are not written in English
5	Studies presenting case studies related to DevOps strategies/guidelines	Studies published before 2011

A checklist (see Table 2) was used as a reference for the quality assessment. Each statement is evaluated based on a three-point numeric scale: yes (two points), partial (one point), and no (0 points). Since there were five questions on the checklist, the maximum number of points any study could be assigned was 10 points; the minimum was 0 points. An article had to receive at least five points to be included in the final primary studies. All items with four or fewer points were excluded from the research for low quality. It is also essential to understand that the assessed article were chosen based on their high relevance to our study. A perfect item could have been excluded if it did not have the desired relevant grade points per the checklist, which was designed to locate highly relevant reviews only. The aim was not to rank the articles depending on their scores; the goal was to exclude articles that contributed little to our research.

Table 2 Sensitivity Analysis Table

Step	Items Checklist	Included (2 point)	Partial (1 point)	Excluded (0 point)
1	Was the problem statement of the study clearly stated?	Yes	1	No
2	Were the initial research gaps and contributions of the study clearly stated?	Yes	1	No
3	Were the results focused on a single relationship and validated?	Yes	1	No
4	Are the limitations and recommendations for future research clearly stated?	Yes	1	No
5	Are the focus of the DevOps approach and its principles clearly defined?	Yes	1	No

Quality assessment using sensitivity analysis was performed in this review by checking whether the selected primary reviews from the previous phase reached the study's minimum quality requirements. Papers that did not meet the stated condition were excluded from the study. The studies that passed this quality assessment (QA) are the final papers considered in the primary research.

2.3 Automated Search of Digital Libraries

The process involved conducting an automated search in the following electronic sources that are relevant to software engineering research:

1. Scopus
2. ACM Digital Library
3. ScienceDirect
4. EBSCOhost
5. SAGE Journals
6. Wiley Online Library
7. ProQuest

Specifically, we searched by title, abstract, and keywords. Papers written in English and published from 2011 to July 2020 were considered. We tried to be as comprehensive as possible.

Consequently, we used a general search string across all explored databases, such as “DevOps guidelines”, “DevOps and guidelines”, “DevOps or guidelines”, and “guideline of DevOps.” We retrieved similar results from databases. We used a single relation in our search string because we wanted to focus on a specific data analysis direction to produce transparent and reliable results.

2.4 The Selection Approach Based on Inclusion and Exclusion Criteria

We started the selection process by reading the title, abstract, and keywords of each study and selecting them according to the criteria in Table 1. Then, we excluded studies that are not primary studies and works that were not peer-reviewed or contained less than four (4) pages (e.g., theses or abstracts). We also excluded books, book titles, and monographs because they usually present mature research work and merge results that are previously presented or consumed by journals or conferences.

We also used a sensitivity analysis approach. After that, we performed a second selection by reading the full text of each included paper. We defined a sensitivity analysis on the item checklist composed of the five criteria in Table 2 and an acceptance or rejection of either “Yes” or “No”, respectively, provided by the individual research assessment scores as shown in the last three (3) columns in Table 2.

The sensitivity analysis aimed to exclude studies of inferior quality per our research focus. This analysis included two stages. In the first stage, each study was read by a randomly selected author assigned to each criterion in Table 2. The acceptance of “Yes” and “No” was provided so that any study that satisfied the sensitivity criteria would be assigned a pass of “Yes” and would be included in the present review. Our “No” sensitivity criteria were excluded if a unique research paper failed. The second stage of the sensitivity analysis was carried out to ensure that high-quality documents were selected from those that passed the first stage. The author, different from the initial one, in the early stage, was assigned to read and assess the task using the same procedure used in the first stage and produce a second sensitivity analysis.

2.5 Searching Based on Snowballing

Snowballing [21] is a search technique typically used to complement automated queries. We adopted backward/forward snowballing to identify additional papers that the initial computerised search might not have captured. In each selected preliminary study, we adapted backward snowballing, its list of references, and forward snowballed its citations in Google Scholar (<http://scholar.google.com/>). We first decided to consider all papers published from 2011–2019 that were not included in the initial search; then, we applied the previously used sensitivity analysis to the other retrieved primary studies.

2.6 Assessing the Research Methodology

Lastly, we performed another literature search to verify all the relevant primary studies. We conducted an automated search over the following electronic databases:

1. Wiley Online Library (<https://onlinelibrary.wiley.com>)
2. Ebsco Discovery Service (<http://eds.b.ebscohost.com/eds/search/>)
3. Springer Link (<https://link.springer.com>)

From this search, we found 18 relevant primary studies, confirming the snowballing approach’s validity, reliability, and completeness.

The search results (Table 3) were filtered, and a duplicate was removed from the first assessment.

Table 3 Search Results with the First Assessment

S/N	Databases Used	Search Results	Databases with First Assessment	
1	ACM Dig Lib	27,058	ACM Dig Lib	117
2	Ebsco Discov Service	1,570	Ebsco Dis S.	83
3	Ebsco Academic Search Complete	18	Ebsco Academic Search Complete	12
4	Google Scholar	3,723	Google Scholar	17
5	Scopus	14	Scopus	3
6	SAGE Journal	01	SAGE Journal	2
7	Science Direct	109	ScienceDirect	23
8	Wiley Oline Lib	89	Wiley OLt	15
9	Springer Link	233	SpringerLink	77
10	ProQuest	1544	ProQuest	09
TOTAL		32,960	TOTAL	358

2.7 Literature Classification Steps

This review classified papers based on the classification criteria and then presented a framework model characterised by different DevOps research areas. The research review derived this framework gradually. Initially, draft schedule topics were generated based on the reading of titles, keywords, and abstracts during the paper-selection stages. This topic comprises six areas with possible branches. We later used these labelled topics to classify the documents while reading the full text and continued to feed new subtopics as needed within each area.

2.7.1 DevOps Concepts

We describe six (6) areas of the DevOps loop flow framework model and their topics based on the DevOps concepts' linkage. All the papers present perspectives on all other relevant DevOps research areas. They address DevOps definitions and methodologies, practice, principles, strategies, benefits, challenges, and software domains concerning the adoption and future research for manageable DevOps implementation.

2.7.2 DevOps Adoption Methodologies

Studies belonging to this category present a methodological approach to DevOps implementation, such as DevOps models and automation.

2.7.3 DevOps Practices

The papers belonging to this area describe DevOps best practices and solutions to some challenges of DevOps adaption. They analyse the definition of DevOps metrics and different strategies for DevOps implementations such as configuration management, continuous testing, continuous monitoring, continuous integration, and continuous deployment/delivery.

2.7.4 DevOps Principles

These studies mention the different stages of DevOps, such as continuous integration (CI), continuous delivery (CD), and continuous deployment (CD). They also comply with the DevOps pipeline showing a specified performance with reliability, automation, measurement, sharing feedback, and fast software release, among other matters.

2.7.5 DevOps Strategies

These studies evaluated different DevOps teams' activities and various management approaches. Silos and targeted results, providing support and quick responses for each group, analysing the process, and quality assurance evaluation of different software delivery. This also involves defining a strategy, understanding requirements, and selecting and testing tools, as shown in Figure 1.

2.7.6 DevOps Benefits

These papers present DevOps benefits using an automated toolset for both development and operation teams to break the silos arising from each specific application domain, including web transaction payments of web applications and mobile applications.

2.7.8 DevOps Challenges

Based on the articles in this category, present DevOps challenges include different team locations or different countries' policies, a lack of a standard methodology, and the lack of a framework for adoption.

3. Results

Finally, the resulting framework serves as a valuable contribution to DevOps research trends during data analysis.

Figure 1 is a continuous DevOps implementation guide that requires best practices with a delivery cycle comprising planning, development, testing, deployment, release, and monitoring. And with active cooperation between different team members, consisting of a series of activities, such as continuous development, continuous automated testing, deployment, continuous monitoring, continuous integration, and continuous delivery. And high-level principles with the acronym "CAMS" (culture, automation, measurement, and sharing). And that includes culture, constant collaboration, and communication.

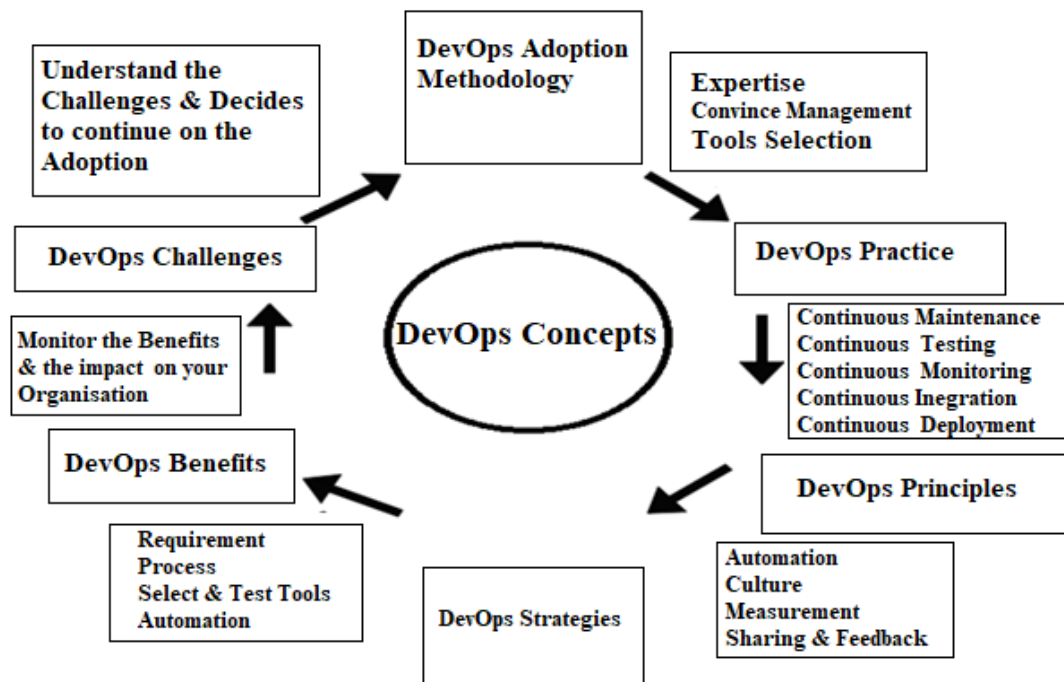


Figure 1 DevOps Loop-Flow Framework Model

Figure 1 is a loop flow framework model that characterises the relevant guides to DevOps implementation and pinpoints the understanding of continuous delivery, which served as a starting point for identifying a common methodology for DevOps adoption in the existing laid-down guidelines.

This research shaped the framework in a looping system because DevOps is a continuous and repeated process along the pipeline.

A further brief explanation of an individual loop flow framework stage is provided presently. This section also reports the number of primary studies selected in each of the previous steps and presents some quantitative analyses of the results, as shown in Table 4.

Table 4 Primary Studies Included, along with the Research Questions Where They Were Mapped

Study Indicators (S)	References	RQs Mapped
S1	[29]	RQ1 & RQ2
S2	[26]	RQ2
S3	[7]	RQ1, RQ2
S4	[19]	RQ2
S5	[13]	RQ1, RQ2
S6	[29]	RQ1, RQ2
S7	[25]	RQ1, RQ2
S8	[22]	RQ1 & RQ2
S9	[23]	RQ1 & RQ2
S10	[23],	RQ1
S11	[15]	RQ1, RQ2
S12	[22]	RQ1, RQ2
S13	[24]	RQ1
S14	[14]	RQ2
S15	[25]	RQ2
S16	[6]	RQ1 & RQ2
S17	[7]	RQ1 & RQ2
S18	[5]	RQ2
S19	[8]	RQ2
S20	[23]	RQ1
S21	[25]	RQ2 & RQ1
S22	[1]	RQ1
S23	[26]	RQ2
S24	[27]	RQ1
S25	[28]	RQ1
S26	[29]	RQ1
S27	[30]	RQ1
S28	[31]	RQ1, RQ2
S29	[32]	RQ1
S30	[30]	RQ2
S31	[33]	RQ2
S32	[34]	RQ2
S33	[35]	RQ2

The data was extracted based on the themes and research questions. Except for DevOps adoption, six (6) themes were identified, with three (3) pieces considered for each research question. Articles were selected based on their impact and relevance to answering the research questions and their strong support from the primary literature. The selected articles are listed in Table 5.

Table 5 Identified Themes and the Primary Studies Supporting Them.

Themes		Study Indicators
DevOps Methodology		
CI/CDs Practice & Collaborations		S8, S13, S14, S15, S16, S17, S19, S23 S25, S28
Strategies and Shared Responsibility		S11, S12, S14, S16, S17, S24, S25
Principles, Domains And guideline		S10, S13, S16, S20, S25, S28, S30, S33
DevOps Adoption		
CI/CD Pipelines		
Monitoring		S12, S19, S20, S23
Automation		S12, S14, S15, S18, S19, S23, S24 S9, S11, S14, S17, S21, S23, S24, S25, S26
DevOps Advantages		
Faster release		
Cycles and		
Low Cost		S11, S12, S13, S20, S21, S23
Higher productivity		
and efficiency		S8, S11, S12, S13, S14
Quality		S11, S12, S13, S20, S27, S28
DevOps Challenges		
Lack of skills and		
Knowledge		S11, S12, S20, S21, S24, S25
Resistance to		S5, S6, S13, S14, S15, S17
Change (Silos)		S2, S5, S6, S17, S19, S20, S23
Un-standardise Tools and High cost		

3.1 Numerical Output

The automated search described in 2.3 returned an initial collection of 358 primary studies. The detailed results of the digital libraries considered in this step are reported in Table 4 and filtered in this initial collection according to the inclusion and exclusion criteria stated in Table 2. Of these, 33 studies passed the two-step quality assessment, as presented in Figure 3.

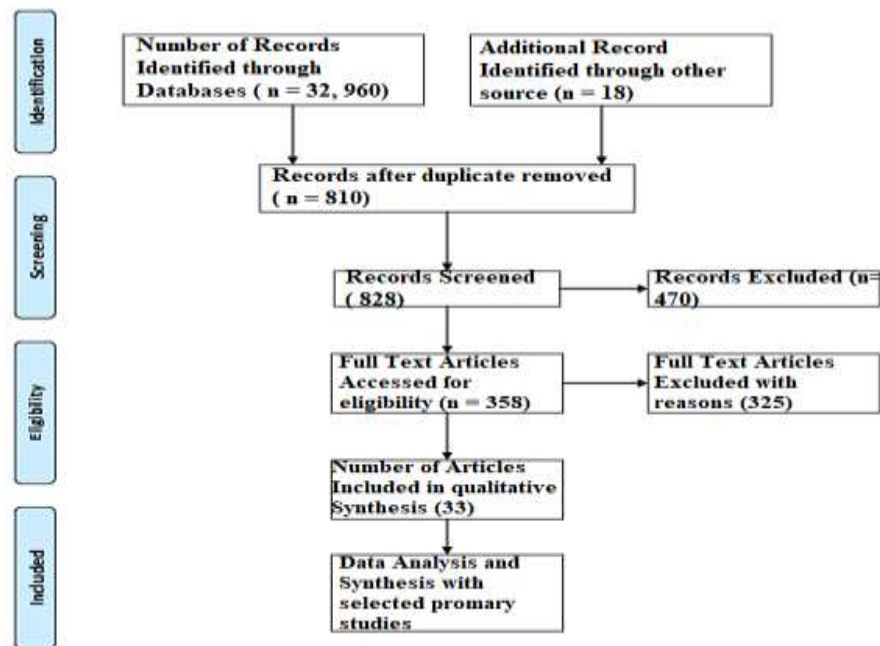


Figure 2 Numerical Search Selection Flow Output Using the PRISMA Model

We launched the query search in July 2020 and applied both forward and backward snowballing iterations (see Section 2.6) on the selected 89 papers. In the snowballing step, 18 new peer-reviewed articles were identified, of which 33 passed the quality assessment selection from the total primary studies. The snowballing and quality assessment were iterated twice consecutively, generating a total of 368 (i.e. 350 + 18) studies. Iteration refers to the second iteration, the third iteration, and the fourth iteration that has not produced new results.

Another automated search was launched after the snowballing process, following the procedure presented in Section 2.3. As illustrated in Table 3, **33 primary studies** were collected from all databases. This result confirmed that our snowballing systems were exhaustive; thus, we proceeded to the analysis and reporting stage.

3.2 Quantitative Analysis

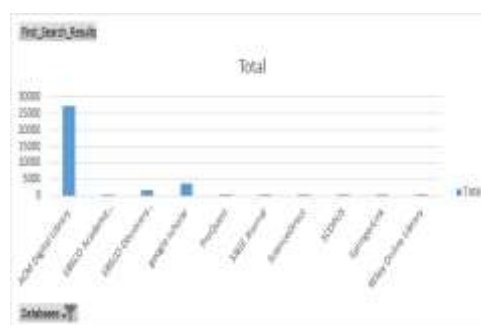


Figure 3, Database-wise Articles Distribution

Figure 3 presents the distribution of articles by the database. This allows us to find any article related to the topic we are researching.

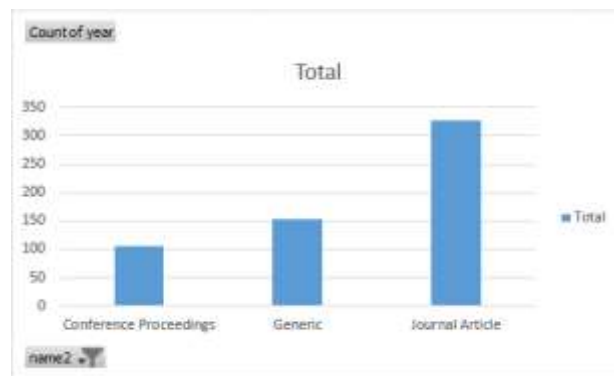


Figure 4 Distribution Based on Categories

Figure 4 represents the number of articles per category.

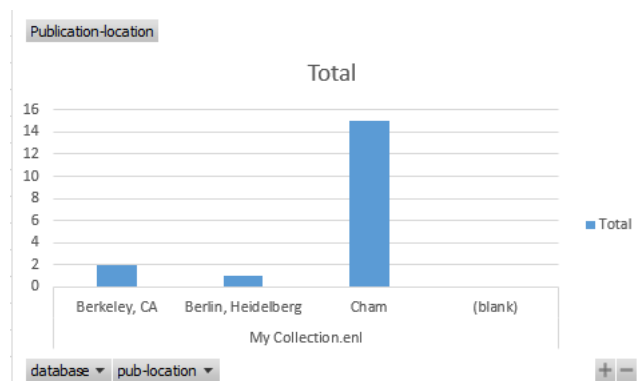


Figure 5 Location-Wise Distribution of Articles

Figure 5 represents the number of articles per location.

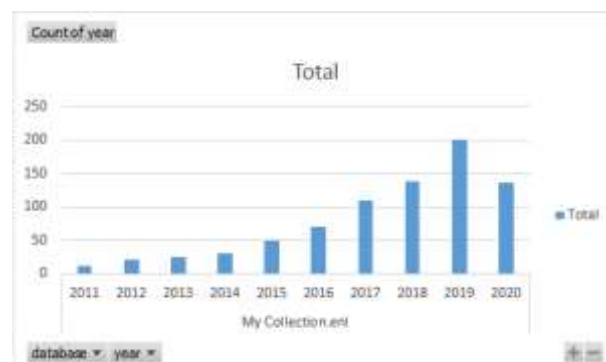


Figure 6 Year-Wise Distribution of Articles

Figure 6 shows the number of articles published each year.

The following sections explain how research questions were answered.

3.2.1 DevOps Guidelines/Steps for Adoption (RQ1)

The adoption of DevOps can face obstacles since there is no straightforward method or framework for this process. Organisations might also want to adopt DevOps in different ways and approach their respective other purposes based on their current DevOps maturity level. This

section will provide some general guidelines and suggestions related to the methods, strategies, and tools that aid the DevOps adoption process based on the selected primary studies.

3.2.1.1 Continuous Integration and Continuous Deployment Pipelines

Studies [S12, S19, S20, S28] have highlighted the main objectives of DevOps adoption. The goals are to achieve continuous delivery and successfully implement and practice continuous integrations in a deployment/delivery pipeline. The pipeline involves automated steps throughout the delivery process, from version control to testing the code to deploying it in production. Each action validates and pushes the code change when necessary and, if approved, sends the code to the next stage. They halt the process if the code change fails to pass pipeline stages, and developers are notified to resolve the related issues.

One study [S28] explained the continuous integration and continuous deployment/delivery stages to fully automate the delivery process, thus reducing the release cycle times, costs, and software development risks. However, the researchers defined the CI/CD pipeline as the last step of the supply chain in the software development process. Automated pipelines can be implemented in many ways using many different tools. Some suggestions presented in [S12] consist of the following steps: unit tests, platform tests, delivery to staging, application acceptance tests, deployment to production, and post-deployment tests, as presented in the pipeline stages. Two studies [S20, S28] showed the same pipeline stages, including building, automated testing, and deployment. Based on the previous reviews [S12, S19, S28], the pipeline was a significant obstacle preventing implementation, which required a wide range of tools and tests. The entire process is monitored; we will discuss the monitoring in the next section.

3.2.1.2 Monitoring

Surveillance and monitoring are vital parts of DevOps and could apply to most aspects of the software development process. In DevOps, the monitoring process is always automated continuously [S12, S14, S15, S18, S24, S28, S29]. Other researchers have claimed that monitoring is paramount as part of the continuous delivery process and is logged continuously aggregated during the coding phase [S15]. Other researchers [S14, S30] divided monitoring into two categories: infrastructure monitoring. The infrastructure enhanced the planning and development processes to increase the value and business results. However, according to a preliminary study [S14], monitoring automation is “the ability to monitor the applications and infrastructure without human intervention.” The monitoring approach was intended to control the system’s functionality and state and indicate to the developers if something in the system had abnormalities. Message alerts can be sent using a chat tool, such as Slack or Hip Chat [S12, S14, S18]. A preliminary study [S18] presented a model for DevOps task categorisation and communication. The model, which monitors automation, is a system actor within the SDLC; workers monitor information captured from the source control system (SCS) and the state and functionality of all other systems within the SDLC. They are also notified if an issue occurs.

Similarly, according to another study [S12], dashboards are the primary service used for monitoring releases, users using the system, and users’ initial country. Some others were also dashboards that enable special teams to monitor their part of the application. Finally, a preliminary study [S19] discussed monitoring to confirm whether a deprecated feature is still in use or whether it can be removed from the application.

3.2.1.3 Tools

The primary studies on DevOps tools [S9, S11, S14, S17, S21, S23, S26, S28, S29] present tools as an essential core factor in considering DevOps adoption. DevOps tools are helpful throughout the development process and within the CI/CD pipeline. These primary studies mainly discussed tools used to control, communicate, monitor, test, and release. Collaboration tools are discussed in two primary studies [S14, S23]. Collaboration tools such as Slack and Hip Chat are typically used for DevOps communication by using the messaging approach to report alarms through notifications via monitoring tools. Preliminary studies [S11, S21, S23, S26] mainly considered git-based tools, such as GitHub and Bitbucket, for version control. Git enables each developer to control their code in a local repository. The developer can then push the changes to a central repository or pull other developers' modifications.

The previous versions of the code are also available when needed if any changes have to be reverted. Monitoring tools, including New Relic, Graylog, and Kinesis, are also discussed in [S11]. These tools help monitor automation and report error signals to teams, as well as graphs and dashboards that represent the monitored data. Based on primary studies [S11, S26], testing can use Selenium or TestNG with the support of Jenkins's automation tool. Selenium can also test for clicks, links, CSS, text, tag names, etc. When testing is completed, tools such as Ansible, Chef, Jenkins, Puppet, and Docker are used for automation, containerisation, delivery, and deployment. These are also some of the tools that enable continuous integration and the CI/CD pipeline [S11, S17, S21, S23, S26].

3.2.1.4 Adoption Strategies

Preliminary studies [S1, S4, S5, S6, S10, S11, S23, S28, S29, S30, S32] further showed that there are still many open questions about how organisations follow guidelines when adopting DevOps. It is encouraging for DevOps practitioners to consider that DevOps adoption requires the support of top management. Furthermore, [S32] explained that it does not happen in the beginning and that a "guerrilla" strategy can take place (i.e., acting outside the company's standard procedures). However, arguments for DevOps adoption can differ between engineers to managers. As such, convincing top management for a successful DevOps adoption and implementation. According to [S29, S30, S32, S33], very close integrations between the management, version control, automation tools, and deployment stages would simplify the process of establishing guidelines and promote compliant development strategies.

3.2.2 DevOps Challenges (RQ2)

Adopting DevOps was clear regarding its benefits, as well as some challenges related to its adoption. Several articles provided an overview of the challenges of adopting DevOps [S8, S9, S11, S12, S13, S20, S22, S24, S26, S27, S28, S63]. These studies have discovered three main themes and challenges organisations face that want to adopt DevOps. The articles lacked expertise, skills, and knowledge of the individual implementing DevOps; strategies for overcoming obstacles to the resistance to change by the relevant stakeholders; and cost implications.

3.2.2.1 Lack of Standard Guidelines and Knowledge Skills

The literature indicates that DevOps lacks a precise method and guidelines or framework for its implementation. The results show that companies were hindered by uneducated personnel regarding DevOps, making it very difficult to adopt. However, the lack of clear guidelines, skills,

and knowledge associated with DevOps is an extensive challenge discussed in many articles related to DevOps challenges [S1, S2, S3, S8, S9, S10, S11, S12, S20, S22, S24, S23, S27]. These articles present an apparent problem with the lack of skills, knowledge, education, and clear guidelines on how to adopt DevOps. In [S11], the researchers discovered some required prerequisite skills and knowledge to develop, test, integrate, and deploy software to productions. In one of their cases, a company adopting DevOps had difficulties related to the technologies and platforms they used in their implementation. Even in other companies whose employees had a high technology and knowledge skillset, the DevOps practitioners had problems with the DevOps approach. In another setting, a primary study [S12] revealed that recruiting the correct personnel with adequate skill sets is paramount. However, a lack of knowledge and required skill sets led to disastrous DevOps adoption and implementation:

“Unqualified personnel having no required skills can slow the DevOps adoption journey because the competencies needed are missing at times of need.”

Four primary studies [S20, S22, S24, and S27] concluded that one of DevOps’ main challenges is the lack of clear guidelines, understanding, and expertise. Finally, the researchers in [S27] established that DevOps adoption guidelines are lacking and can lead to delays in the software release cycle (Figure 1). [S10] provided similar guidelines for medical software that suggest the following: (1) fostering a community to change the culture to establish collaborative channels and merge traditional mindsets with agile (they are complementary, not in conflict); (2) co-creating a standard DevOps IT platform while also supporting legal development; (3) improving, iterating, and gathering insights and key performance metrics; and (4) growing processes over time, over-communicating, telling stories, sharing success, and showing data.

3.2.2.2 Obstacles and Resistance to Change

One of the challenges to adopting DevOps is the lack of valuable strategies, leading relevant stakeholders to resist change. Some DevOps-relevant stakeholders might resist DevOps adoption for fear of losing their jobs or their refusal to agree with the collaborative phenomenon. Different teams might have different orientation skills and perform other tasks that might negatively impact cooperation [S12, S20]. According to [S22], there are conflicting goals between development and operations teams. Developers tend to want new features and bugs fixed to be released rapidly. Simultaneously, the operations team wants to keep the releases to a minimum to ensure smooth operational activities to preserve the system’s stability and reliability. From another perspective, senior management is resisting DevOps adoption. The literature provides some views on the problems encountered due to senior management. For example, according to [S29], while the benefits of adopting DevOps are unclear, top management may resist DevOps adoption by questioning DevOps implementation’s capability and wisdom.

Tools for DevOps automation are relatively expensive and need proper management, which may deter managers from understanding DevOps’ value. DevOps adoption is sometimes affected at a particular hierarchical level and by an inflexible management style [S13, S22, S24]. The researchers in [S24] discussed the lack of productivity at the beginning of the adoption process, which may send a wrong signal as a problem for senior management. However, senior management can delay adopting DevOps due to such productivity issues.

3.2.2.3 Methodological Shortfalls & the Cost of Tools and Automation

The primary studies uncovered several challenges: the lack of standard methodological approach and tools for any particular organisation, cost implications, and insufficient knowledge of the automation process. Preliminary studies [S1, S5, S6, S7, S9, S12, S13, S24, S26, S27, S28]

presented a lack of tools as a frequent challenge for adopting DevOps. Due to a lack of standard DevOps practices, relevant stakeholders have difficulties finding the correct tools and methods for their organisations and understanding how to use them reliably for their continuous implementation [S9, S12, S27]. The initial arrangement of the tools, experimentation, and deciding which tool to use are considered negative time factors [S12]. DevOps tooling's relatively high cost is also considered a negative factor, and management may have difficulties justifying their investment in the DevOps journey. One preliminary study [S27] pointed out that a lack of standardisation between process and tools hinders DevOps adoption. Therefore, devices can be a liability if handled poorly. The preliminary study [S28] stated:

“The evidence of this study shows that the asset can become a liability if the control of the resources is managed wrongly. Respondents mentioned that the failure of resources control could lead to resource overhead during the integration of the source codes when all source codes have been deployed too often. As a result, this will jeopardise the Continuous Delivery Pipeline. In some cases, the automated test failed because the production environment is different and very complex to be executed with the automated test”.

Considering the statement above, the continuous deployment of features can be too complicated, leading them to fail continuous delivery pipeline tests.

3.2.3 DevOps Adoption Advantages (RQ3)

There are many benefits of adopting DevOps, according to the literature analysis. Similarly, in the literature, we identified eight (8) articles that mentioned these benefits [S1, S2, S3, S11, S12, S13, S21, S27]. These articles provide similar conclusions about DevOps' benefits. However, organisations should know DevOps benefits for those who want to adopt the DevOps journey.

3.2.3.1 Frequent Release Cycle

The frequent release cycle describes a method that begins with completing code and delivers the code to production. Concerning the DevOps phenomenon, the release time is often short, ranging from a few hours to a few weeks, possibly involving several daily releases. The six articles above agree that faster release cycles benefit from adopting DevOps. [S8, S20].

A case study [S11] discussed a shorter release cycle time. It was stated that *“An improved speed in the delivery of software changes was the most commonly perceived benefit of DevOps”*. This was the perspective presented by four out of five case companies in their study. Release times were reduced to a few days from several months. Other articles [S12, S13, S27] also mentioned that DevOps reduces the release cycles and suggested that release cycles are an essential benefit of DevOps adoption. Similarly, the survey outcome results of [S28, S10] (as presented in Figure 1 above) indicate the following:

“The majority of the respondents highlighted the advantage of the DevOps activities that they reduce the software cycle time according to what they believe DevOps practice can achieve.”

This assertion was also consistent with the other literature highlighting DevOps' benefits; one can confirm that one of DevOps' key advantages is a “shorter and faster release cycle.”

3.2.3.2 High Throughput (Output)

Throughput explains several ways to measure the effectiveness of a process or system in converting inputs into outputs. High productivity is determined by dividing the result by the costs incurred or resources spent. With software development, productivity is measured by backend code. The number of code lines can compare the number of regulations to another factor, such as

the time required to implement the lines of principles, measured as lines of code per unit time. Some DevOps practices increase productivity, such as automation, collaboration, communication, and improved strategies implemented by all concerned teams. The primary study provided some support for improved productivity when adopting DevOps. Prior studies [S11, S12, S28] indicated that better strategies in breaking the silos of communication and less bureaucracy improved the productivity of development and operations teams. However, the development and operations team must improve their communication to enhance collaboration, which would, in turn, increase productivity. When adopting DevOps, large portions of (if not all) bottlenecks are eliminated, significantly growing different teams' productivity. [S13] claimed that continuous integration and feedback increase the productivity of development teams. Productivity is also enhanced by sharing knowledge across various groups to manage the shared understanding and track the knowledge needs throughout the development life cycle [S9].

3.2.3.3 Quality

This research considered the quality of teams and the quality of codes and applications. The literature on DevOps' benefits unanimously indicates that code quality and application quality are essential benefits of DevOps [S11, S12, S13, S20, S28]. Three primary studies [S11, S12, S28] conducted case studies and analysed companies that adopted DevOps. They concluded that quality increases whenever DevOps practices are implemented. They also ascertained that they tend to produce higher quality codes, increasing assigning developers responsibility for their respective codes.

Similarly, developers are more confident that their code will pass the tests and deploy into the production environment with more minor incremented releases. Two primary studies [S13 and S20] further discuss the QA of DevOps and conclude that DevOps can drive QA by enhancing communication and feedback loops, as presented in Figure 2. A study [S21] reported that most companies feel that DevOps improves production quality and decreases the risks. The automation of processes, such as continuous testing and continuous deployment, substantially impacted the code's quality. Shorter and faster release cycles enabled deployment in smaller increments, increasing production quality since the risks and quality might be easy to control.

3.3 Threats to Validity

Validity and reliability reveal threats to internal-external, dependability, transferability, and empirical research study validity.

Internal validity has to do with the confidence of the analysed and reported results, and the study considered framework definition and adoption. We have benefited from the description of the classification framework and its usage for different research classifications. It is a threat to these studies. However, we make the classification data available to allow the research community to evaluate our results' validity.

Another threat to internal validity is framework inclusiveness. The reviewed classification framework might not include all areas and topics characterising DevOps guidelines research. If the risk is eliminated, the framework will be derived sequentially and incrementally. As we have previously explained, starting from the first draft framework obtained from reading only the title, abstract, and keywords, new subtopics have been added within each area after reading the whole paper.

External validity threats are derived from potential issues that prevent the absolute generalisation of the results. Only a subset of papers concerning DevOps guidelines was targeted in our study (i.e., we only considered papers published within a period of nine (9) years, from 2011 to July 2020). However, the results may not be well represented or represent the overall research in the field. We acknowledged that the risk of this is low because DevOps guidelines are a new research topic; as such, very soon, other investigations will likely include the most relevant research fields. One relevant threat was an automated search performed in July 2020. The area of DevOps is growing fast, and future studies would likely find more recent papers not included in the searched or used databases. This threat has been resolved by forwarding the snowballing approach, which gave us the chance to include many additional papers citing the primary studies with publication years of 2011–2020.

Finally, construct validity comprises threats related to the correspondence of measures managed to the related properties. Based on our approach, the following threats were highlighted.

We identified the primary studies. We initially defined a search string to identify prior studies in our approach. Some or most of the time, different search strings might have produced different results. However, this is a general threat to all systematic review approaches. We tried to overcome this issue first by redefining a very available search string to be as widespread and comprehensive as possible. The next threat to our primary study's identification was the consideration of digital databases. At the beginning of our search, we used the seven most popular databases for software engineering. Still, we thought searching on different sources might have yielded different results. This risk was considerably eliminated using various loop iterations of backward and forward snowballing as a search procedure to complement our search of originally selected databases. Lastly, we performed some verification tasks by launching another second automated search on three (3) more databases, and the results confirmed the validity of the search results previously conducted. We have seen it as very unlikely to miss all the relevant articles.

Selection of primary studies: The exclusion or inclusion of an article by reading the title, abstract, and keywords according to the inclusion and exclusion criteria was initially specified. Then, the whole paper was read based on a two-step quality assessment approach. There is a possibility that some reports were missed due to the defined inclusion/exclusion criteria or to the illustrated quality assessment checklist of the above selection procedure. Meanwhile, in representing such a selection procedure, we had strictly followed the guidelines for systematic reviews in software engineering [36] and the selection procedures followed in similar studies [21].

3.4 Summary of the Findings

The review identified DevOps benefits, challenges, and common DevOps guidelines for adoption. Our discussion of the findings highlights answers to the research questions in line with the six areas of the classification framework model proposed earlier. The main results are summarised in Table 6. The research started by identifying some general challenges to all aspects of DevOps adoption, which have a broad impact and, thus, return across the loop flow areas. The common challenges include the following:

3.4.1 Lack of A Standard Methodology for DevOps Adoption

There was no transparent scientific approach to determining the DevOps capabilities. This challenge was described in studies concerning methodological evaluation and portability across

different domains, such as medical, finance, embedded system domains, and other highly regulated environments and policy variations.

3.4.2 Technology Evolution Constraints

There are constraints concerning the continuous and rapid evolution of DevOps tools' technology. Transversally impacts all areas and the loop flow (refer to Figure 1). Practising activities will continuously face novel challenges. It is always difficult to adapt to new technology constraints and conditions, especially considering the ongoing deployment problems in some working environments. Applying the DevOps methodology is not straightforward since it is highly regulated and often involves legal issues.

3.4.3 Cost of Tools

The cost of tools ranges from the large dimensions and high complexity of DevOps systems to their different configurations. The high prices of these tools heavily impact their automation, continuous integration, continuous delivery, and constant testing.

3.4.4 Policy Restrictions (different teams' locations)

Concerning different policy capabilities and cross-platform DevOps tools, the automation for provisioning and de-provisioning resources and services is another challenge that directly affects the continuous deployment policy perspective. It requires specific approach guidelines for different organisations.

3.4.5 Security Issues

Security issues are of crucial concern in other domains and DevOps. Such issues arise in DevOps and entail even more difficulties than the traditional software development approach. The response to security issues was instantiated as "DevSecOps."

Table 6 Summary of the main findings

DevOps Challenges	The primary studies indicated that a lack of standardisation between process and tools significantly hinders DevOps adoption. There is no precise method, framework, or guidelines for its adoption.	Looking clearly at primary studies related to the DevOps computing discipline, there was no transparent scientific approach to determining DevOps' capabilities regarding process, practice, tools, and implementation.	Another frequent challenge discovered from the primary studies was the lack of standard methodological approach and tools for any particular organisation, cost implications, and insufficient knowledge of the automation process.
DevOps Adoption Benefits	Primary studies show that improving speed	The primary studies indicated that the software's	The literature on the benefits of DevOps

	in the delivery of software changes was the most commonly perceived benefit of DevOps.	quality increases whenever DevOps best practices are implemented and adopted successfully.	unanimously showed that code quality and application quality are essential benefits of DevOps.
Classification Adoption Guide	Documentation of projects and planning Prepared expertise in the organisation Ability to automate the process Building an effective DevOps culture across all teams Implementation of a code management strategy Ability to implement the principles (continuous integration/development/deployment/delivery)		

4. Conclusions

The primary studies reviewed in the present work provide adequate information about the benefits and challenges of DevOps implementation. The guidelines to be followed were precisely identified based on the different modes of a report from the primary studies. The review identified DevOps benefits, challenges, and common adoption guidelines. The primary studies considered in this review show that DevOps adoption is currently being rolled out in software development worldwide. We examined the benefits, challenges, and guidelines of DevOps adoption. Throughout this process, we explored various publications on DevOps' benefits, challenges, and guidelines using multiple databases, including ACM Digital Library, Google Scholar, and ScienceDirect. The present review also revealed that the automation, sharing, and collaboration of the DevOps process are closely related to benefits and successes in software development based on common guidelines. We also found a direct link between DevOps implementation without guidelines and various challenges, such as insufficient knowledge of the automation process. There was no transparent scientific approach to determining DevOps' capabilities regarding processes, practices, tools, and implementation. The current review also revealed that DevOps is the fastest available response software development process. It is also essential that DevOps is accompanied by common guidelines. However, DevOps contributes to ensuring continuous software delivery and shortening time to market. Our research shows that research has focused primarily on process automation, continuous delivery, and integration. Further research is needed in many DevOps areas (e.g. measurement, the development of performance metrics of different stages, practice, and culture) to ensure the appropriateness of software reliability and quality factors such as efficiency, usability, and portability. Finally, other survey research questions categorised by different roles can be posed and answered by DevOps practitioners.

Acknowledgement

The authors would like to acknowledge the support from the School of Computer Sciences, USM for the publication of this paper.

References

- [1] E. Dornenburg, "The Path to DevOps," *IEEE Software*, vol. 35, no. 5, pp. 71–75, 2018, DOI: 10.1109/MS.2018.290110337.
- [2] R. Jabbari, N. bin Ali, K. Petersen, and B. Tanveer, "What is DevOps? A systematic mapping study on definitions and practices," *ACM International Conference Proceeding Series*, vol. 24-May-201, May 2016, DOI: 10.1145/2962695.2962707.
- [3] R. Jabbari, N. bin Ali, K. Petersen, and B. Tanveer, "What is DevOps? A systematic mapping study on definitions and practices," in *ACM International Conference Proceeding Series*, 2016, vol. 24-May-201, pp. 1–11. DOI: 10.1145/2962695.2962707.
- [4] M. Daneva and R. Bolscher, "What we know about software architecture styles in continuous delivery and devops?," in *Communications in Computer and Information Science*, vol. 1250 CCIS, 2020, pp. 26–39. DOI: 10.1007/978-3-030-52991-8_2.
- [5] G. Österberg, "A Systematic Literature Review on DevOps and its Definitions, Adoptions, Benefits and Challenges," Masters thesis, Åbo Akademi University, 2018. [Online]. Available: <https://www.doria.fi/handle/10024/177223>
- [6] K. Maroukian and S. R. Gulliver, "Leading DevOps Practice and Principle Adoption," no. June, pp. 41–56, 2020, DOI: 10.5121/csit.2020.100504.
- [7] M. Artač, T. Borovšak, E. di Nitto, M. Guerriero, and D. A. Tamburri, "Model-Driven continuous deployment for quality devops," *QUDOS 2016 - Proceedings of the 2nd International Workshop on Quality-Aware DevOps, co-located with ISSTA 2016*, pp. 40–41, 2016, DOI: 10.1145/2945408.2945417.
- [8] M. Dlugi, A. Brunnert, and H. Krcmar, "Model-based performance evaluations in continuous delivery pipelines," *1st International Workshop on Quality-Aware DevOps, QUDOS 2015 - Proceedings*, pp. 25–26, 2015, DOI: 10.1145/2804371.2804376.
- [9] L. Riungu-Kalliosaari, S. Mäkinen, L. E. Lwakatare, J. Tiihonen, and T. Männistö, "DevOps adoption benefits and challenges in practice: A case study," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2016, vol. 10027 LNCS, pp. 590–597. DOI: 10.1007/978-3-319-49094-6_44.
- [10] M. Zufahmi Toh, S. Sahibuddin, and M. N. Mahrin, "Adoption issues in DevOps from the perspective of continuous delivery pipeline," in *ACM International Conference Proceeding Series*, 2019, vol. Part F1479, pp. 173–177. DOI: 10.1145/3316615.3316619.
- [11] R. Jabbari, N. bin Ali, K. Petersen, and B. Tanveer, "Towards a benefits dependency network for DevOps based on a systematic literature review," *Journal of Software: Evolution and Process*, vol. 30, no. 11, Nov. 2018, DOI: 10.1002/smr.1957.
- [12] F. Erich, C. Amrit, and M. Daneva, "Cooperation between information system development and operations: A literature review," *International Symposium on Empirical Software Engineering and Measurement*, vol. 26, no. June 2002, p. 2652598, 2014, DOI: 10.1145/2652524.2652598.
- [13] R. Jabbari, N. bin Ali, K. Petersen, and B. Tanveer, "What is DevOps? A systematic mapping study on definitions and practices," *ACM International Conference Proceeding Series*, vol. 24-May-201, 2016, DOI: 10.1145/2962695.2962707.
- [14] M. Airaj, "Enable cloud DevOps approach for industry and higher education," *Concurrency and Computation: Practice and Experience*, vol. 29, no. 5, Mar. 2017, DOI: 10.1002/cpe.3937.

- [15] S. Kolb and C. Rock, "Unified Cloud Application Management," *Proceedings - 2016 IEEE World Congress on Services, SERVICES 2016*, pp. 1–8, 2016, DOI: 10.1109/SERVICES.2016.7.
- [16] L. E. Lwakatare *et al.*, "DevOps in practice: A multiple case study of five companies," *Information and Software Technology*, vol. 114, pp. 217–230, 2019, DOI: 10.1016/j.infsof.2019.06.010.
- [17] J. Díaz, D. López-Fernández, J. Pérez, and Á. González-Prieto, "Why are many businesses instilling a DevOps culture into their organization?," *Empirical Software Engineering*, vol. 26, no. 2, 2021, DOI: 10.1007/s10664-020-09919-3.
- [18] V. Desai and H. A. Dinesh, "INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY Efficient Reputation-based Cyber Attack Detection Mechanism for Big Data Environment," 2022, DOI: 10.17485/IJST/v15i13.2102.
- [19] R. Jabbari, N. bin Ali, K. Petersen, and B. Tanveer, "Towards a benefits dependency network for DevOps based on a systematic literature review," *Journal of Software: Evolution and Process*, vol. 30, no. 11, Nov. 2018, DOI: 10.1002/smr.1957.
- [20] P. Brereton, B. A. Kitchenham, D. Budgen, M. Turner, and M. Khalil, "Lessons from applying the systematic literature review process within the software engineering domain," *Journal of Systems and Software*, vol. 80, no. 4, pp. 571–583, 2007, DOI: 10.1016/j.jss.2006.07.009.
- [21] C. Wohlin, "Guidelines for snowballing in systematic literature studies and a replication in software engineering," *ACM International Conference Proceeding Series*, 2014, DOI: 10.1145/2601248.2601268.
- [22] K. Meixner, D. Winkler, and S. Biffl, "Towards Combined Process & Tool Variability Management in Software Testing," 2019, DOI: 10.1145/3302333.3302339.
- [23] R. K. Gupta, M. Venkatachalapathy, and F. K. Jeberla, "Challenges in Adopting Continuous Delivery and DevOps in a Globally Distributed Product Team: A Case Study of a Healthcare Organization," *Proceedings - 2019 ACM/IEEE 14th International Conference on Global Software Engineering, ICGSE 2019*, pp. 30–34, 2019, DOI: 10.1109/ICGSE.2019.00020.
- [24] M. Zufahmi Toh, S. Sahibuddin, and M. N. Mahrin, "Adoption issues in DevOps from the perspective of continuous delivery pipeline," in *ACM International Conference Proceeding Series*, 2019, vol. Part F1479, pp. 173–177. DOI: 10.1145/3316615.3316619.
- [25] J. A. Morales, H. Yasar, and A. Volkman, "Implementing DevOps practices in highly regulated environments," *ACM International Conference Proceeding Series*, vol. Part F1477, pp. 1–9, 2018, DOI: 10.1145/3234152.3234188.
- [26] M. Fazal-Baqaie, B. Güldali, and S. Oberthür, "Towards DevOps in multi-provider projects," *CEUR Workshop Proceedings*, vol. 1806, pp. 18–21, 2017.
- [27] R. Vallon, B. J. da Silva Estácio, R. Prikladnicki, and T. Grechenig, "Systematic literature review on agile practices in global software development," *Information and Software Technology*, vol. 96, pp. 161–180, Apr. 2018, DOI: 10.1016/j.infsof.2017.12.004.
- [28] L. E. Lwakatare *et al.*, "DevOps in practice: A multiple case study of five companies," *Information and Software Technology*, vol. 114, pp. 217–230, Oct. 2019, DOI: 10.1016/j.infsof.2019.06.010.
- [29] A. F. Nogueira, J. C. B. Ribeiro, M. Zenha-Rela, and A. Craske, "Improving la redoute's CI/CD pipeline and devops processes by applying machine learning techniques," in *Proceedings - 2018 International Conference on the Quality of Information and*

- Communications Technology, QUATIC 2018*, Dec. 2018, pp. 282–286. DOI: 10.1109/QUATIC.2018.00050.
- [30] S. Mandepudi, "Communication Challenges in DevOps & Mitigation Strategies," 2019, [Online]. Available: <http://www.diva-portal.org/smash/record.jsf?pid=diva2:1368400>
 - [31] M. F. Lie, M. Sánchez-gordón, and R. Colomo-palacios, "DevOps in an ISO 13485 Regulated Environment : A Multivocal Literature Review," 2020.
 - [32] M. Senapathi, J. Buchan, and H. Osman, "DevOps capabilities, practices, and challenges: Insights from a case study," *ACM International Conference Proceeding Series*, vol. Part F1377, no. July 2019, 2018, DOI: 10.1145/3210459.3210465.
 - [33] T. Laukkarinen, K. Kuusinen, and T. Mikkonen, "Regulated software meets DevOps," *Information and Software Technology*, vol. 97, pp. 176–178, May 2018, DOI: 10.1016/j.infsof.2018.01.011.
 - [34] T. Laukkarinen, K. Kuusinen, and T. Mikkonen, "DevOps in regulated software development: Case medical devices," *Proceedings - 2017 IEEE/ACM 39th International Conference on Software Engineering: New Ideas and Emerging Results Track, ICSE-NIER 2017*, pp. 15–18, 2017, DOI: 10.1109/ICSE-NIER.2017.20.
 - [35] L. E. Lwakatare *et al.*, "Towards DevOps in the embedded systems domain: Why is it so hard?," *Proceedings of the Annual Hawaii International Conference on System Sciences*, vol. 2016-March, no. February 2020, pp. 5437–5446, 2016, DOI: 10.1109/HICSS.2016.671.
 - [36] J. Pérez, J. Díaz, J. Garcia-Martin, and B. Tabuenca, "Systematic literature reviews in software engineering—enhancement of the study selection process using Cohen's Kappa statistic," *Journal of Systems and Software*, vol. 168, Oct. 2020, DOI: 10.1016/j.jss.2020.110657

A REVIEW: CHARACTERIZATION OF ALUMINIUM-METAL MATRIX COMPOSITE REINFORCED WITH ZIRCONIUM DIBORIDE

Muhammad Azim Suhaimi¹
Mohd Muzamir Mahat¹
Sabrina M. Yahaya²
Rosmamuhamadani Ramli^{1*}

¹Material Sciences and Technology Programme, Faculty of Applied Sciences, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia

²Applied Chemistry Programme, Faculty of Applied Sciences, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia

Abstract: Composite materials with aluminium matrices are used in thermal management, defence, aerospace, automotive and aviation. As the prices of useful items decreased, their usefulness increased. Aluminium is reinforced with other metals, non-metals and ceramics to give it the required physical and mechanical properties such as high tensile strength, great hardness and corrosion resistance. Aluminium-metal matrix composites are known for their exceptional physical and mechanical properties and performance. This work also focuses on the fabrication and characterization of zirconium diboride reinforced aluminium-metal matrix composite. Zirconium diboride particles were used as reinforcement material to fabricate aluminium metal matrix composites by the stir casting method. The microstructure, hardness and tensile strength of aluminum-metal matrix composites and its corrosion behaviour were investigated. The results showed that the inclusion of zirconium diboride particles increased the hardness and tensile strength of aluminum-metal matrix composites. Field emission scanning electron microscopy and x-ray diffraction analysis are used to investigate the microstructural properties of aluminum-metal matrix composites reinforced with zirconium diboride particles. Field emission scanning electron microscopy images showed a uniform distribution of zirconium diboride particles in the aluminum-metal matrix composites matrix. X-ray diffraction analysis showed the formation of an intermetallic compound of aluminium and zirconium diboride. From the review that have been made, we can conclude that the inclusion of zirconium diboride particles increased the mechanical properties and corrosion of aluminium-metal matrix composites.

Keywords: Aluminium-Metal Matrix Composite, Zirconium Diboride, Mechanical Properties, Corrosion Behaviour

1. Introduction

Researchers and manufacturers have been keeping an eye on metal matrix composites (MMCs) for some time to investigate their physical and mechanical properties and performance. MMCs are composites that differ by their base metal such as aluminium (Al), copper (Cu), or titanium (Ti), depending on the reinforcement, the type of phase (fibres, particles, whiskers), and the manufacturing processes such as powder metallurgy, diffusion bonding, infiltration, and stir casting [1]. These processes allow different combinations of composite properties such as tensile strength, hardness, and corrosion resistance depending on the application. For example,

cast iron and aluminium-silicon (Al-Si) have long been produced in foundries. The combination of these two materials results in a composite material that has superior properties that cannot be achieved with monolithic alloys [2]. The properties can also be changed and modified based on the volume, shape and sizes of the phase diagram, which contains all the details of the information about the metal composites.

Aluminium matrix composites (AMC) are used because they have undeniably better properties, namely high mechanical strength and good thermal conductivity. The main objective in the production of AMC is to create a metal that has low mass and density while increasing stiffness. For this reason, Al alloys are one of the materials chosen as the basis for the use of composites, as they are suitable for most large manufacturers, being light and strong. As general, the study shows that the addition of Cu to pure Al increases hardness, reduces grain size and improves the mechanical properties of the metal. Aluminium-copper-zirconium diboride (Al-Cu-ZrB₂) is a MMCs produced by an in-situ reaction that improves mechanical properties. Zirconium diboride (ZrB₂) causes the Al-Cu to increase tensile strength and hardness. Depending on the type of reinforcement, AMC leads to different structures and process-related properties. Surappa and Mirle [3] show that the different processes produce different properties, sizes and shapes. The four process types such as particle reinforced AMCs, whisker or short fibre reinforced AMCs, continuous fibre reinforced AMCs, and monofilament reinforced AMCs are the processes used to produce AMC materials.

ZrB₂, a low-density, ultra-high temperature ceramic material, is a good candidate for aerospace applications. It has relatively high thermal and electrical conductivity. ZrB₂ serves as a grain refiner and greatly improves the mechanical properties of composite materials. ZrB₂ reinforcement has been used with many matrices including Al, Cu, magnesium (Mg) and their alloys [4]. Sivakumar *et al.*, [5] found that the size of agglomerates decreased significantly when ZrB₂ was added to Al in larger amounts. This means that the size of agglomerates in the Al of the particles in the metal composite decreases, the ZrB₂ particles occupy spaces that were not filled in the base metal Al. It is also mentioned that pure Al has rather low hardness, but the mechanical properties of Al-ZrB₂ composites were greatly improved. Al-20wt.% ZrB₂ had a very high yield strength (529 MPa), compressive strength (630 MPa), and compressive elongation of 19.25%. The use of ZrB₂ as reinforcement increases the strength and improves the wear resistance.

2. Casting Technique

Casting is one of the various techniques used to produce MMCs. In the case of Al, conventional casting is used to produce alumina, silicon, and tungsten carbide particles, as well as short-fibre and whisker alumina, silicon carbide, and carbon. In this case, three types of casting processes were analysed: In-situ, ex-situ and stir casting processes. Both processes are used to produce zirconium diboride enriched MMC.

2.1 In-situ technique

Recently, in-situ methods have been developed for the assembly of Al-based MMC, which can lead to better interfacial adhesion and better mechanical properties. However, the composites treated by these techniques exhibit a number of problems, including thermodynamic imbalance of reinforcement within the matrix, brittle matrix-reinforcement interface with inhomogeneous distribution of reinforcing elements, and loss of good mechanical properties at elevated

temperatures. To overcome these problems, researchers have developed in situ fabricated composites.

The literature addresses Al composites fabricated in-situ with various reinforcements in the form of tiny particles to evaluate the morphology and performance of MMCs. In situ Al-based composites were scaled up to improve mechanical characterization through dispersion strengthening and grain refinement achieved by the presence of particles in the melt during the solidification process [6]. In addition, AA2024/ZrB₂ AMCs were prepared by direct melt reaction of potassium fluorozirconate (K₂ZrF₆) and potassium tetrafluoroborate (KBF₄), and ZrB₂ micro-sized particles were found to improve the ductility of the composite.

2.2 Ex-Situ Technique

Ex-situ techniques are methods for studying the properties and microstructure of materials outside their natural environment. These techniques are widely used in materials science and engineering to study the properties of Al-MMC materials. A common ex-situ technique to study Al-MMCs is transmission electron microscopy (TEM). In a study by Li *et al.*, [7], the microstructure and mechanical properties of aluminium-silica carbide (Al-SiC) composites were investigated with TEM. They found that the silica carbide (SiC) particles were well distributed in the Al matrix and that the Al-SiC MMCs exhibited better mechanical properties compared to pure Al.

Another ex-situ technique commonly used to study Al-MMCs is scanning electron microscopy (SEM). According to Zhang *et al.*, [8], the microstructure of powder metallurgically prepared Al-SiC MMCs was investigated by SEM. They found that the SiC particles were well dispersed in the Al matrix and that the Al-SiC MMCs exhibited better mechanical properties compared to pure Al.

Mechanical Properties

3.1 Tensile properties

The effect of ZrB₂ reinforcement on the tensile properties of Al-MMCs was studied by <<<<<. They found the specimens were prepared with different weight percentages of ZrB₂ (0, 2, 4, , and 8 wt.%) and tensile strength tests were conducted using a universal tester. The results showed that the addition of ZrB₂ significantly improved the tensile properties of the composite, with the highest tensile strength value of 224 MPa found for the composite reinforced with 8% ZrB₂, while pure Al had a tensile strength value of only 95 MPa. The improvement in tensile properties was attributed to the ZrB₂ particles, which served as reinforcing phases, increasing the overall strength of the material. The improved tensile properties were also attributed to the strong bonding between the Al matrix and the ZrB₂ particles. The results are similar to those of previous studies on Al-based composites reinforced with ZrB₂, but are somewhat lower than the results of a recent work by Pandey *et al.*, [9] due to differences in the fabrication and testing procedures. The work provides useful insights into the potential applications for ZrB₂ reinforced AMMC.

The comparable reinforcement used both of TiB₂ and ZrB₂ to Al-Cu also was investigated by Anis *et al.*, [10]. In their research, they found that Al-Cu with 3ZrB₂ displayed higher strength compared to with 3 wt.%TiB₂ which showed that the stronger interfacial bonding between the Al dendrites and ZrB₂ particulates.

3.2 Hardness Properties

Anis *et al.*, [10] also stated that the hardness value of ZrB₂ displayed higher compared to TiB₂ when reinforced to Al-Cu alloy. This statement was supported by the research that have been done by Filho *et al.*, [11]. According to [11], mechanical hardness testing is essential for the evaluation of the mechanical properties of metallic materials, since the hardness values can be used to determine the wear resistance of the material as well as the approximate values of its ductility, yield stress and other important properties. They also mentioned the definition of mechanical hardness testing, which is used to evaluate some of the most important mechanical properties of materials.

In their study, the effect of ZrB₂ as reinforcement on the hardness parameters of the fabricated Al-MMCs is investigated. The vickers hardness test was used to determine the hardness of composite samples with different amounts of ZrB₂ reinforcement (0, 2, 4, 6, and 8 wt.%). The results show that the Al-MMCs reinforced with ZrB₂ is much harder than pure Al. The composite with the highest hardness, 142 Hv, was found with 8 wt.% ZrB₂ reinforcement. The presence of ZrB₂ particles in the composite matrix serves as hard reinforcing phases and increases the overall hardness of the material. Moreover, the bonding between the Al matrix and the ZrB₂ particles leads to the improved hardness properties. The increase in hardness is exactly proportional to the volume fraction of reinforcement. The results of this research are similar to those of previous studies that investigated Al-based composites reinforced with ZrB₂; however, differences in preparation and testing methods may cause slight variations in the results. This research supports the use of ZrB₂ reinforcement to improve the mechanical properties of Al-MMCs.

3.3 Corrosive Behaviour

This work describes the corrosion properties of Al-MMCs reinforced with ZrB₂. The microstructure of Al-MMCs samples prepared by stir casting with different weight percentages of ZrB₂ particles was investigated by scanning electron microscopy (SEM). In a 3.5 wt% sodium chloride (NaCl) solution, the corrosion behaviour of the samples was investigated using the potentiodynamic polarisation method. The study showed that the addition of ZrB₂ particles to the Al matrix significantly increased the corrosion resistance of the Al-MMCs, with the samples reinforced with 15 wt.% ZrB₂ exhibiting the highest corrosion resistance. The improvement in corrosion resistance is due to the formation of a protective oxide layer on the surface of the samples, which acts as a barrier to the diffusion of corrosive ions and reduces the corrosion rate. The results are consistent with the findings of Liu *et al.*, [12], who found that the corrosion resistance of Al-MMCs reinforced with ZrB₂ increases with increasing reinforcement content. However, the results contradict the findings of Khan *et al.*, [13], who found that the addition of ZrB₂ particles to an Al matrix slightly decreased the corrosion resistance of Al-MMCs. The discrepancy is due to the different study designs and situations.

With reference to a study by Rosmamuhamadani *et al.*, [6], corrosion tests were performed on each alloy in its as-cast condition. Potentiostatic polarisation experiments were performed using a Radiometer Analytical model PG 100 Potentiostat/Galvanostat and VoltaLab software. This method also allows to vary the main results related to the corrosion mechanisms, the rust rate and even the corrosion susceptibility of a particular reinforced alloy in specific settings. The Al-MMC samples were prepared by stir casting method with different weight percentages (0, 5, 10 and 15 wt.%) of ZrB₂ particles. The corrosion behaviour of the samples was evaluated by the potentiodynamic polarisation method in a 3.5 wt% NaCl solution.

Mustafa *et al.*, [14], in their research have investigate the effect of ZrB_2 to Al-Cu alloy. Corrosion potential diminished with rising ZrB_2 content in the alloys. The corrosion current and corrosion rate reveals an exemplary value at 6 wt. % ZrB_2 . Inclusion of ZrB_2 has an effective alteration on corrosion barrier in Al-Cu. This action might be associated to beneficial influence of ZrB_2 inclusion on grain size refinement. The fine-grained materials have more conveniences behaviour with account to corrosion together with oxidation have been established [8]. From Figure 1 below, Al-Cu alloy with 6wt.% ZrB_2 displayed the optimum corrosion barrier compare to other compositions. Value of corrosion rate at 3 wt.% ZrB_2 was $2.85 \times 10^{-3} \text{ mm y}^{-1}$, then decreased to $0.36 \times 10^{-3} \text{ mm y}^{-1}$ for Al-Cu with 6wt.% ZrB_2 . The susceptibility of the Al-Cu alloys towards corrosion decreases in the order of:

$\text{Al-Cu-6wt.\%ZrB}_2 > \text{Al-Cu-3ZrB}_2 > \text{Al-Cu}$.

The least value of corrosion rate demonstrated the alloy has a favourable trait to endure corrosion.

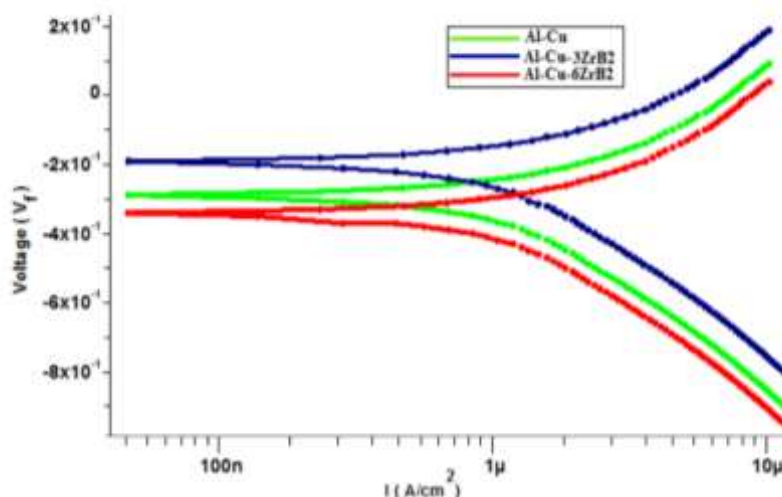


FIGURE 1: Potential-Dynamic Polarization Curves on Al-Cu Alloy Reinforced by ZrB_2 [14]

3.4. Characterization

3.4.1 Field Emission Scanning Electron Microscopy

Recent research has characterised the microstructure and mechanical properties of Al-MMCs with the addition of ZrB_2 using FESEM. Wang *et al.*, [15] and Li *et al.*, [7] discovered that the addition of well-dispersed ZrB_2 particles into the Al matrix can significantly improve the mechanical properties of Al-MMCs, such as hardness and tensile strength, compared to pure Al. However, Zhang *et al.*, [16] found that the Al-MMCs adversely affect the mechanical properties when the ZrB_2 particles are insufficiently dispersed in the matrix. To maximise the potential improvement in mechanical properties, good dispersion of the reinforcement particles throughout the composite must be achieved. These results indicate that the manufacturing process must be carefully controlled to maximise the dispersion of ZrB_2 particles and achieve the necessary improvements in mechanical properties.

Another study by Wang *et al.*, [17] used FESEM to investigate the effects of heat treatment on the microstructure of Al-SiC MMCs and found that SiC particles were more uniformly distributed after heat treatment. They suggested that this uniform distribution could lead to an improvement in the mechanical properties of the composite. In addition to microstructure analysis, FESEM can also be used to analyse the composition of Al-MMCs. For example, in a study by Li *et al.*, [18], FESEM was used with energy dispersive X-ray spectroscopy (EDS) to investigate the elemental composition of Al-B₄C MMCs. They found that the B₄C particles were well distributed in the Al matrix and that the Al-B₄C MMCs exhibited better mechanical properties compared to pure Al.

In a research conducted by Yadav and Dixit [19], they use FESEM to perform a comparative study on the erosion-corrosion of aluminium-silicon (Al-Si) alloys (AA336) and their composites AA336-7% SiC and AA336-7% TiB₂ in basic, acidic and marine atmospheres. The microstructure in the inter dendritic region of Al shows the formation of Si in the eutectic state. Figure 2 (a) shows a classifiable microstructure of the Al-Si alloy (AA336) in the as-cast state. Figure 2 (b) shows the microstructure of the composite reinforced with SiC particles. SEM shows the homogeneous distribution of SiC.

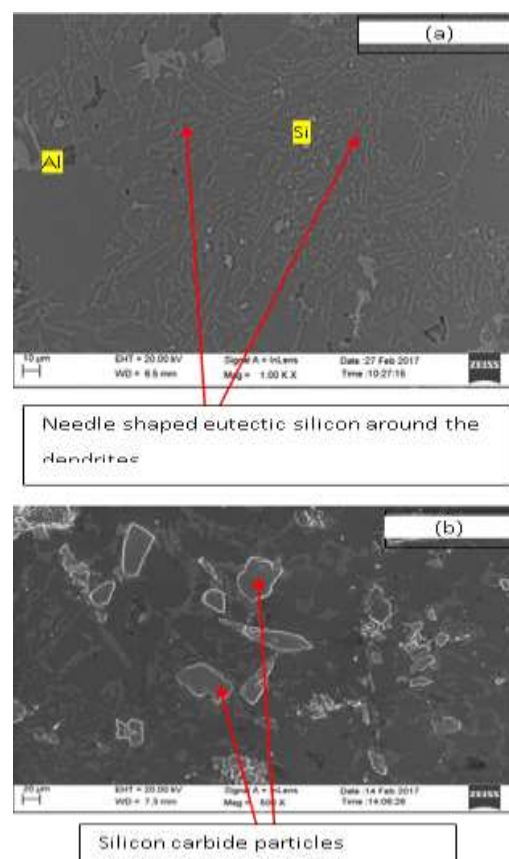


Figure 2: Micrograph on the Polishes Surfaces of (A) Alloy, B) SiC Composites [19]

3.4.2 X-Ray Diffraction

Several studies have investigated the crystal structure of Al-MMCs reinforced with ZrB₂ by x-ray diffraction (XRD). To ensure good dispersion of the reinforcement particles and improve the mechanical properties of the composite, the manufacturing process must be closely

controlled. Using XRD analysis, Chen *et al.*, [20], Liu *et al.*, [21] and Wang *et al.*, [22] have investigated the crystal structure and mechanical properties of Al-MMCs reinforced with ZrB₂. Chen *et al.*, [20] found that the addition of ZrB₂ led to the formation of a new phase, particularly zirconium aluminide (Al₃Zr), and improved the mechanical properties, such as higher hardness and tensile strength. Liu *et al.* [21] also found that the addition of ZrB₂ led to the formation of Al₃Zr and improved the mechanical properties such as strength and toughness. Wang *et al.*, [22] found that the Al-MMCs exhibited improved mechanical properties and the formation of Al₃Zr and aluminium diboride (AlB₂).

The addition of ZrB₂ to the Al matrix leads to the formation of new phases, such as Al₃Zr and AlB₂, and improves the mechanical properties of the composite, such as hardness, tensile strength, and toughness, compared to pure Al. However, the microstructure and properties of Al-MMCs are strongly influenced by the fabrication method and the amount of reinforcement.

In the context of Al-MMC, XRD can be used to study the crystal structure of the Al matrix, the reinforcement particles, and the interface between the two. For example, in a study by Li *et al.*, [18], the crystal structure of Al-SiC MMCs was investigated using XRD and found that the Al matrix has a face-centred cubic (FCC) crystal structure, while the SiC particles have a hexagonal crystal structure. The also found that the Al-SiC MMCs have better mechanical properties than pure Al due to the reinforcing SiC particles. In another study by Zhang *et al.*, [23], XRD was used to investigate the effect of heat treatment on the crystal structure of aluminium boron carbide (Al-B₄C) MMCs. The authors found that the B₄C particles had a hexagonal crystal structure and that the Al-B₄C MMCs exhibited better mechanical properties after heat treatment due to the formation of a stronger interface between the Al matrix and the boron carbide (B₄C) particles.

4. Conclusion

The use of Al-MMCs reinforced with ZrB₂ has attracted considerable attention in recent years due to their potential applications in various fields such as aerospace, automotive, and biomedical. The incorporation of ZrB₂ into the Al matrix results in improved mechanical properties such as higher hardness, tensile strength, strength, and toughness compared to pure Al. Characterization of Al-MMCs by XRD analysis is widely used to study the crystal structure and phase formation of the composite.

It is important to note that the fabrication process and the amount of added reinforcement can greatly affect the final microstructure and properties of the Al-MMCs. Therefore, it is crucial to carefully control the fabrication process to achieve good dispersion of the reinforcement particles and optimize the mechanical properties of the composite.

Acknowledgement

The author would like thanks to Petronas Research Sdn. Bhd. (UTVSB/CS/P 20221001006) for financial support. Special thanks also to dedicate to the Faculty of Applied Sciences, and Faculty of Mechanical Engineering, UiTM Shah Alam, Selangor for facilities and guidance.

References

- [1] Mortensen, A., and Llorca, J. Metal matrix composites. Annual review of materials research, 40 (2010): 243-270. <https://doi.org/10.1146/annurev-matsci-070909-104511>
- [2] Rohatgi, P. K. Metal matrix composites. *Defence science journal*, 43(4) (1993); 323.

- [https://www.scirp.org/\(S\(i43dyn45teexjx455qlt3d2q\)\)/reference/ReferencesPapers.aspx?ReferenceID=642811](https://www.scirp.org/(S(i43dyn45teexjx455qlt3d2q))/reference/ReferencesPapers.aspx?ReferenceID=642811)
- [3] Surappa, M. K. Aluminium matrix composites: Challenges and opportunities. *Sadhana* 28, 92003): 319-334. <https://doi.org/10.1007/BF02717141>
 - [4] Kumar, V., Kumar, S., Kumar, S., and Gupta, M. “Development of Aluminium Metal Matrix Composite Reinforced with ZrB₂ Particles: A Review. *Materials Today*” *Proceedings*, 16, (2019): 2328-2335. doi: 10.1016/j.matpr.2019.05.167
 - [5] Sivakumar, S. Brahma, R.G., and Koteswararao, V.R. “Influence of ZrB₂ hard ceramic reinforcement on mechanical and wear properties of aluminium”. *Ceram. Int.* 45, (2019): 7055–7070. dOI:[10.1016/j.ceramint.2018.12.208](https://doi.org/10.1016/j.ceramint.2018.12.208)
 - [6] Rosmamuhamadani, R., Sulaiman S. Hanim M. A., Ismail M. I. S. and Talari M. K. “Mechanical and Microstructure Characterization of Aluminium – Copper (Al-Cu) Reinforced with in-Situ Titanium Diboride (TiB₂)”, *Key Engineering Materials*, vol. 673 (2016): 117 – 125. doi:[10.14419/ijet.v7i4.14.27691](https://doi.org/10.14419/ijet.v7i4.14.27691)
 - [7] Li, X., Chen, S., and Wang, D. “Microstructure and properties of Al-SiC metal matrix composites fabricated by powder metallurgy”. *Journal of Alloys and Compounds*, 816, (2020). 151835. <https://doi.org/10.3390/met8040253>
 - [8] Zhang, Y., Zhu, L., and Zhu, J. “Microstructure and properties of Al-SiC metal matrix composites fabricated by powder metallurgy”. *Materials Science and Engineering A*, 722 (2019): 191-197. <https://doi.org/10.3390/met8040253>
 - [9] Pandey, K. K., and Tripathy, S. “Characterization of Aluminium Metal Matrix Composite reinforced with Zirconium Diboride using powder metallurgy method”. *Journal of Materials Science and Engineering*, 6(3), (2021): 137-143.
 - [10] Anis, A.L, R. Rosmamuhamadani, Darham, W, Zakaria, A, and Talari, M.K. “Microstructure and Mechanical properties investigation of in-situ TiB₂ and ZrB₂ reinforced Al-4Cu composites”. *IOP Conference Series: Materials Sciences and Engineering* 114. (2016). doi:[10.1088/1757-899X/114/1/012120](https://doi.org/10.1088/1757-899X/114/1/012120).
 - [11] Filho, A. F., Ferreira, I. L., Da Costa, C. E. F., and Gomes, A. S. “The role of hardness tests in the study of wear of metallic materials”. *Materials Research*, 13(2), (2010): 235-242. doi: [10.1590/S1516-14392010000200015](https://doi.org/10.1590/S1516-14392010000200015)
 - [12] L Liu, B., Gao, F., Wang, Y., Zhou, W., and Yan, M. “Corrosion behavior of aluminum matrix composites reinforced with ZrB₂ particles” *Journal of Materials Research and Technology*, 9(2), (2020): 2119-2126. doi: [10.1016/j.jmrt.2019.11.051](https://doi.org/10.1016/j.jmrt.2019.11.051)
 - [13] Khan, Z. A., Rahman, M. M., and Ahmad, Z. “Corrosion behavior of Al–ZrB₂ composites fabricated by powder metallurgy route”. *Journal of Materials Research and Technology*, 8(1), (2019): 603-611. doi:[10.1016/j.jmrt.2018.12.018](https://doi.org/10.1016/j.jmrt.2018.12.018).
 - [14] N.H. Mustafa, M.M. Mahat, S.M. Yahaya and R. Rosmamuhamadani. Corrosion behaviour of *in-situ* Zirconium Diboride (ZrB₂) reinforced by Aluminium-Copper (Al-Cu) alloy metal matrix composite. *Journal of Physics: Conference Series*, Volume 1349. doi: [10.1088/1742-6596/1349/1/012083](https://doi.org/10.1088/1742-6596/1349/1/012083)
 - [15] Wang, J., Zhang, Y., and Zhu, L. “Microstructure and mechanical properties of Al-ZrB₂ metal matrix composites fabricated by powder metallurgy”. *Journal of Alloys and Compounds*, 785, (2019): 721-726.
 - [16] Zhang, Y., Zhu, L., and Zhu, J. “Microstructure and mechanical properties of Al-ZrB₂ metal matrix composites fabricated by powder metallurgy”. *Materials Science and Engineering A*, 789 (2021): 1-5.
 - [17] Wang, J., Zhang, Y., and Zhu, L. “Microstructure and properties of Al-SiC metal matrix composites fabricated by powder metallurgy with heat treatment”. *Journal of Alloys and*

- Compounds*, 789 (2019): 735-741.
- [18] Li, X., Chen, S., and Wang, D. "Microstructure and properties of Al-B₄C metal matrix composites fabricated by powder metallurgy". *Journal of Alloys and Compounds*, 853 (2021): 156958.
- [19] Yadav, P. K., and Dixit, G. "Erosive-corrosive wear of aluminium-silicon matrix (AA336) and SiCP/TiB₂p Ceramic Composites". *Silicon*, 11(3) (2018): 1649–1660. <https://doi.org/10.1007/s12633-018-9981-2>.
- [20] Chen, L., Wang, J., and Liu, Y. "Microstructure and properties of Al-ZrB₂ metal matrix composites fabricated by powder metallurgy". *Journal of Alloys and Compounds*, 753 (2018): 2-6.
- [21] Liu, Y., Chen, L., and Wang, J. "Microstructure and properties of Al-ZrB₂ metal matrix composites fabricated by powder metallurgy". *Journal of Materials Science and Engineering A* (2019): 765, 1-5.
- [22] Wang, X., Chen, L., and Wang, J. "Microstructure and properties of Al-ZrB₂ metal matrix composites fabricated by powder metallurgy". *Journal of Materials Science*, 45(16) (2020): 5894-5901.
- [23] Zhang, Y., Zhu, L., and Zhu, J. "Microstructure and properties of Al-SiC metal matrix composites fabricated by powder metallurgy". *Materials Science and Engineering A*, 722 (2018): 191-197.

THE RELATIONSHIP BETWEEN BUILDING CONDITION SPECIFICALLY TO INDOOR AIR QUALITY (IAQ) AND SICK BUILDING SYNDROME (SBS) SYMPTOMS AMONG WARD STAFF AT HOSPITAL PAKAR SULTANAH FATIMAH

Mohd Arif Rosli^{1*}
Muhamad Shahril Mohd Abdullah^{1*}
Mohd Syafiq Syazwan Mustafa^{1*}
Azian Hariri²
Amir Abdullah Muhamad Damanhuri³
Nur Ain Ebas⁴
Nor Mohd Razif Noraini

¹Faculty of Engineering Technology, Universiti Tun Hussein Onn Malaysia (UTHM)

²Faculty of Mechanical and Manufacturing Engineering, Universiti Tun Hussein Onn Malaysia (UTHM)

³Faculty of Mechanical and Manufacturing Engineering Technology, Universiti Teknikal Malaysia Melaka (UTeM)

⁴Faculty of Civil Engineering and Built Environment, Universiti Tun Hussein Onn Malaysia (UTHM)

Abstract: *This study aims to investigate the relationship between building condition, particularly indoor air quality, and Sick Building Syndrome (SBS) symptoms among ward staff at Hospital Pakar Sultanah Fatimah. The study will employ a cross-sectional design and a sample of ward staff will be recruited to participate in the study. SBS symptoms will be assessed using a standardized questionnaire. Statistical analysis will be conducted to examine the relationship between building condition, indoor air quality, and SBS symptoms. The findings of this study will contribute to the understanding of the impact of building condition on indoor air quality and the health of healthcare workers. The calculated U (63) is greater than the critical U (53), indicating that H_0 cannot be rejected. Therefore, we accept H_0 which states that the present symptoms and building conditions are equal, and we can conclude that there is a relationship between building conditions and present symptoms among the ward staff.*

Keywords: *Sick Building Syndrome (SBS), Indoor Air Quality (IAQ), Contaminants, Health*

1. Introduction

Recently, Malaysia and generally all over the world have been shocked by the issue of covid 19 pandemic. Rahardiman *et al.*, [1] found several areas within the hospital during a pandemic, such as the emergency room and treatment rooms, where the concentrations of bacteria and fungi were slightly elevated. The authors suggest that these findings may be due to the increased use of personal protective equipment and cleaning agents during the pandemic, which may have affected the hospital's ventilation system and indoor air quality. Besides, Fonseca *et al.*, [2] reviewed 64 articles published between 2015 and 2020, and they found that indoor air quality (IAQ) in healthcare units is a complex issue that requires a comprehensive approach. The study reveals that the most common indoor air pollutants in healthcare units are biological

contaminants, such as bacteria, fungi, and viruses, as well as chemical pollutants, such as volatile organic compounds (VOCs) and particulate matter. The sources of these pollutants include patients, medical equipment, cleaning agents, and outdoor air pollution.

Dikmen & Gültekin [3] discussed in their research paper the importance of maintaining good indoor air quality in hospital buildings, and the challenges that hospitals face in achieving sustainable IAQ. There are several sources of indoor air pollution in hospital buildings as described by the author, including biological pollutants such as mold, bacteria, and viruses that can cause respiratory infections and allergic reactions, chemical pollutants such as volatile organic compounds (VOCs) that are released from cleaning agents, disinfectants, and medical equipment, outdoor pollutants such as particulate matter, nitrogen oxides, and sulfur dioxide that can enter the building through ventilation systems or open windows and the occupational pollutants such as latex particles, anesthetic gases, and drugs that can pose a risk to healthcare workers who are exposed to them on a regular basis. The finding can be supported by reviewing the previous research from [4-7] which discuss detailed reviews of the sources of indoor air pollutants in hospital environments, along with recommendations for monitoring and mitigating their impact on indoor air quality.

The health risks associated with poor indoor air quality in hospital buildings are significant. Patients with weakened immune systems, such as those in intensive care units, are particularly vulnerable to the health effects of poor IAQ. Healthcare workers who are exposed to indoor air pollutants on a regular basis may also be at risk for occupational illnesses. Therefore, it is important to maintain good indoor air quality in hospitals to ensure a safe and healthy environment for everyone. The gap for this research are specifically focusing on the impact of building conditions on indoor air quality and sick building syndrome (SBS) symptoms among hospital staff. Most of the previous studies have been conducted in office buildings, and there is a need to investigate the relationship between building conditions, indoor air quality, and SBS symptoms in hospital settings. This study significantly contribute to the development of a better understanding of the relationship between building conditions, indoor air quality, and SBS symptoms among ward staff in a hospital setting in Malaysia. The study's findings will provide important information for hospital administrators and building managers to improve the indoor environment of their facilities, thereby reducing SBS symptoms and improving staff productivity and well-being. The aim of this study is to investigate the relationship between building conditions, specifically indoor air quality IAQ and SBS symptom and identify the factors that influence the perception of building conditions and indoor air quality among ward staff at Hospital Pakar Sultanah Fatimah.

2. IAQ and SBS Symptoms at Hospital Condition

According to the International Labour Organization (ILO), Sick Building Syndrome (SBS) is defined as a phenomenon that occurs when 20% of employees report experiencing symptoms associated with their workplace, particularly in relation to air quality [8]. Sick Building Syndrome (SBS) refers to a range of nonspecific subjective health symptoms that can include itchy eyes, skin rashes, and nasal allergy symptoms, as well as more general symptoms such as fatigue, aches and pains, sensitivity to odours, and difficulty in concentration [9]. The impact on health increases with longer and more intense exposure, leading to a range of symptoms from minor eye irritations, coughing, wheezing, and allergic rhinitis to more severe respiratory symptoms, decreased lung and heart function, tuberculosis, cardiovascular diseases, and even premature death[10-15].

The study by Zainal *et al.*, [16] found that higher levels of VOCs and PM were associated with increased reports of SBS symptoms. The concentration of indoor air pollutants exceeded the recommended levels reported a high prevalence of SBS symptoms, including eye irritation, headaches, and dry or itchy skin [17]. Table 1 provides a summary of the latest research studies on indoor air quality (IAQ) in hospital conditions. The table discusses the IAQ parameters that were studied, the IAQ problems identified, the sick building syndrome (SBS) symptoms reported, and the methods used to control IAQ levels.

Table 1: Several Previous Research Study About the IAQ in Hospital condition

Previous research	IAQ Parameter discussed	SBS Symptom	Method Control IAQ
[18]	Temperature, relative humidity (RH), carbon dioxide (CO ₂), volatile organic compounds (VOCs), particulate matter (PM _{2.5} and PM ₁₀)	Headaches Fatigue Eye irritation	Improving ventilation, reducing indoor pollutant sources by increasing air filtration
[19]	Temperature, RH, CO ₂ , VOCs, particulate matter (PM _{2.5} and PM ₁₀) formaldehyde, nitrogen dioxide (NO ₂), ozone (O ₃), carbon monoxide (CO)	Headaches Fatigue Eye irritation	Improving ventilation Increasing air filtration Regular maintenance and cleaning of HVAC systems
[20]	Temperature, RH, CO ₂ , VOCs, particulate matter (PM _{2.5} and PM ₁₀) formaldehyde, NO ₂ , O ₃ , CO, exchange rate, air velocity	Headaches Fatigue Eye irritation Respiratory symptoms	Improving ventilation Increasing air filtration Regular cleaning and maintenance of HVAC systems Using low-emitting materials and furnishings
[21]	CO ₂ , particulate matter (PM ₁₀ and PM _{2.5}), VOCs, formaldehyde, NO ₂ , O ₃ , CO, temperature, RH, air exchange rate	Headaches Fatigue Eye and nose irritation Respiratory symptoms	Improving ventilation Increasing air filtration Regular maintenance and cleaning of HVAC systems Using low-emitting materials and furnishings
[22]	CO ₂ , CO, formaldehyde, VOCs, total fungi count, total bacterial count, temperature, RH	Eye irritation Nasal irritation Dryness or irritation of the throat Fatigue Headaches Dizziness	Increase ventilation rates and use high-efficiency air filters Conduct regular maintenance of HVAC systems, increase the distance between indoor sources of pollution and building occupants
[23]	Temperature, RH, CO ₂ , CO, TVOC, particulate matter (PM ₁₀),	Eye irritation Nose irritation	Implement adequate ventilation systems, monitor and maintain

	total fungal count, total bacterial count	Throat irritation Fatigue Headaches Dizziness	HVAC systems, limit exposure to indoor pollutants, proper disposal of hazardous waste
[1]	Microbiological (bacteria, fungi) pollutants	-	Use of ultraviolet germicidal irradiation (UVGI) and high-efficiency particulate air (HEPA) filters, and regular disinfection of surfaces and air handling units.
[24]	Microbiological pollutants	-	Adequate ventilation, air filtration, and regular cleaning and disinfection
[25]	microbiological pollutants	-	QuEChER method for air microbiological monitoring
[2]	Temperature, RH, CO ₂ particulate matter (PM), VOCs, bacteria, fungi, and other microorganisms.	Headaches, fatigue, eye and throat irritation, and respiratory problems	Proper ventilation and air filtration, regular maintenance of HVAC systems, use of low-emitting materials, and appropriate cleaning

According to reviewing the previous paper as listed in Table 1, there are several common IAQ problems such as poor or inadequate ventilation, inadequate air filtration, high level of indoor pollutants, and elevated indoor air pollutants. Previous research has provided numerous references discussing the health effects, specifically SBS symptoms, that result from exposure to air pollutants. Some examples of references and respected references are Carbon dioxide (CO₂) [26][27], Carbon monoxide (CO) [28], Nitrogen Dioxide (NO₂)[29], Formaldehyde (HCHO) [30][31][32], Total Volatile Organic Compound (TVOC)[33], Respirable Particulate (PM_{2.5}) [34], Ozone (O₃)[35], Total Bacteria count [36][37], Total Fungal Count [38][39]. The primary problem of microbiological contamination arises when the extent of contamination exceeds a particular limit, which is considered ordinary for a specific environment [40].

The listed Indoor air quality (IAQ) parameters discussed in the studies are crucial to understanding and improving indoor air quality. Temperature, humidity, carbon dioxide, volatile organic compounds, and particulate matter are all important factors that can impact indoor air quality and occupant health. Monitoring and controlling these parameters is essential for creating a healthy indoor environment. The IAQ problems identified in the studies are also significant. Poor ventilation, high levels of indoor pollutants, and inadequate air filtration can all contribute to IAQ issues and result in adverse health effects for building occupants. By identifying these problems, researchers can develop solutions to address them and improve IAQ. The SBS symptoms reported by occupants are also important to consider. Headaches, fatigue, and eye irritation are just a few of the symptoms that can result from poor IAQ. These symptoms can lead to decreased productivity and overall discomfort for building occupants.

Finally, the methods used to control IAQ levels are crucial for improving indoor air quality. Improving ventilation, increasing air filtration, and reducing indoor pollutant sources are just a few of the methods that can be used to improve IAQ. By implementing these methods, building owners and managers can create a healthier and more comfortable indoor environment for occupants.

2.1 Establish Guideline Related to IAQ

When delving into the parameters of Indoor Air Quality (IAQ), examining the health effects and exploring common methods to control IAQ levels, it is imperative to take into account established safety and health guidelines. The guidelines that should be referenced are as follows:

- i. The World Health Organization's (WHO) guidelines for Indoor Air Quality (IAQ) regarding Selected Pollutants not only furnish information on the health effects, sources, and levels of exposure, but also provide practical and efficacious recommendations for mitigating and controlling exposure to these deleterious pollutants [41].
- ii. USA OSHA - Another valuable resource for information on hazardous chemicals present in occupational settings is the USA OSHA Occupational Chemical Database. Comprising over 1,000 chemical compounds, this database details their health effects and physical properties and offers practical guidance on safe handling practices, making it an indispensable tool for safeguarding the well-being of workers in diverse industries [42].
- iii. The USA National Institute for Occupational Safety and Health (NIOSH) Pocket Guide to Chemical Hazards is an invaluable resource for those who work with hazardous chemicals. As a quick-reference tool, it provides workers and other stakeholders with critical information about the potential hazards associated with various chemicals and guidelines for their safe handling in the workplace. By promoting safe chemical management practices, the Pocket Guide can effectively prevent workplace injuries and illnesses, making it an indispensable asset for maintaining a healthy and productive work environment [43].
- iv. The Malaysian Department of Occupational Safety and Health (DOSH) Industry Code of Practice on Air Quality is a comprehensive guide for maintaining optimal indoor air quality in the workplace. Offering invaluable guidance on ventilation system design, installation, and maintenance, as well as practical recommendations for minimizing the use of harmful chemicals and pollutants indoors, this code of practice is an indispensable resource for preventing workplace injuries and illnesses caused by exposure to indoor air contaminants [44].

All of the guidelines mentioned above are critical resources for maintaining and safeguarding the health and well-being of workers and other individuals in indoor environments.

3. Methodology

3.1 Study Design

A cross-sectional comparative study was conducted at two selected buildings (ward). These buildings were chosen are ward 16 and ward 18, Hospital Pakar Sultanah Fatimah, also known as Sultanah Fatimah Specialist Hospital. The selection criteria for buildings are based on a long history of occupancy and the age of the buildings, which dates back to 1900, making them

more than 100 years old. The study population comprised the respondents determined 58 staffs consist of nurses, administrative and doctors who are worked at the ward.

3.2 Questionnaire Survey

The study population was given a set of structured questionnaires that were adapted from ICOP 2010. These questionnaires included information on the respondents' socio-demographics, health status, and symptoms of SBS, as well as the amount of time they spent in the building each week, their job position, and the quality of the air inside the buildings. To be classified as having SBS, respondents needed to have at least one symptom that occurred at least once a week and have reported experiencing symptoms for at least 3 days per week in the last four weeks [9][45]. Additionally, respondents needed to have reported an improvement in their symptoms when they were away from their place of work.

3.3 Case Study Location

The Sultanah Fatimah Specialist Hospital, also known as the Government Dispensary, was first established in Jalan Petri, Muar in around 1900. However, due to town planning in 1918, the hospital was relocated to Jalan Salleh. Since the 1920s, the hospital, formerly known as Muar Hospital, has undergone significant development [46]. On October 13, 2003, His Highness the Sultanah Johor officially declared the hospital's name as the Sultanah Fatimah Specialist Hospital. It is currently the third-ranked hospital in Johor, following Sultanah Aminah Hospital and Sultan Ismail Hospital. The hospital also serves as a reference hospital for the northern zone of Johor, which includes Batu Pahat, Tangkak, and Segamat as shown in Figure 1. Additionally, the hospital holds the important role of being a teaching hospital. The hospital's master plan for development began in 1990, and there are plans to build an ambulatory Care Centre Block which will provide sophisticated medical equipment facilities to cater to the specialized services offered by the hospital. However, the construction of this block is still pending.



Figure 1: The location of HPSH (adopted from HPSF main page Website)

4. Results and Discussion

4.1 Building Characteristics

An investigation was conducted to identify the unique characteristics and condition of each building. It was found that both ward buildings were constructed approximately 100 years ago. Additionally, all buildings are situated in a medium-sized town that is in close proximity to heavy traffic and moderately near an industrial area. The buildings themselves are made of either concrete or bricks, with slanting tile roofs, and they all have mechanical ventilation air conditioning systems installed. Smoking is strictly prohibited in these buildings, and daily cleaning procedures involve mopping with water and detergent. Notably, no major renovation activities have been conducted on these buildings in recent decades. During an interview session with the Hospital Administration, which included the Director, Dr. Muhammad Yusof Sibert, the hospital engineer, Mrs. Rogayah and the ward staff, a significant amount of information was gathered about Wards 16 and 18. This included details on the number of staff responsible for these wards, the number of beds, and the annual patient admission rates.

4.2 Respondents Characteristics

The survey results revealed that a significant proportion of respondents exhibited symptoms compatible with SBS. However, due to limitations in the number of available respondents, the validity of the cross-sectional study might be affected. Moreover, selection bias could have occurred because of unequal respondent rates, which may have affected the prevalence of SBS symptoms. A total of 60 questionnaires were distributed, and the response rate was 95%. Socio-demographic data were obtained from respondents in their respective buildings and are presented in Table 2. The proportion of female respondents was higher than that of males. The age of the respondents was 46.6% (25-39), 52.6% (40-55), and 1.8% (> 55). All respondents spent the same number of hours per week in the buildings, which was 40 hours or 8 hours per day.

Table 2: Characteristics of respondents

Building	Ward 16 & 18
N = 60	
Response rate:	95%
Sex distribution:	
Male	6.3%
Female	94.7%
Age of respondents:	46.6% (25-39)
	52.6% (40-55)
	1.8% (>55)
Working time per week (hours):	All respondents 40h/week @ 8h/day

4.3 Workplace Environment

Based on the responses provided by the respondents, it appears that there are several issues related to the room conditions and indoor air quality in the hospital ward. One of the most reported issues is varying room temperature, with 73.7% of respondents indicating that they have experienced this problem as shown in Table 3.

Table 3: Percentage of respondents in workplace environment

Workplace condition	Sample (s) = 57 (%)	
	Yes	No
Draught	20 (35.1)	37(64.9)
Room temperature too high	27(47.4)	30(52.6)
Varying room temperature	42(73.7)	15(26.3)
Room temperature too low	30(52.6)	27(47.4)
Stuffy “bad” air	19(33.3)	38(66.7)
Dry air	17(29.8)	40(70.2)
Unpleasant odour	21(36.8)	36(63.2)
Static electricity causes shocks	4(7)	53(93)
Passive smoking	19(33.3)	38(66.7)
Noise	19(33.3)	38(66.7)
Light that is dim or cause glare and/or reflections	15(26.3)	42(73.7)
Dust and dirt	16(28.1)	42(71.9)

This suggests that there may be an issue with the heating, ventilation, and air conditioning (HVAC) system in the ward that needs to be addressed. Another issue that was reported by a significant number of respondents is room temperature being too low (52.6%) or too high (47.4%). This could have negative impacts on both the patients and the healthcare workers, as excessively warm or cold temperatures can lead to discomfort and potentially even health problems. Additionally, a significant proportion of respondents reported experiencing stuffy or "bad" air (33.3%) and unpleasant odors (36.8%), which could suggest that there are issues with ventilation or cleanliness in the ward. Other reported issues, such as noise (33.3%), dust and dirt (28.1%), and lighting (26.3%), could also have negative impacts on the comfort and well-being of patients and healthcare workers.

Overall, these survey results suggest that there may be several areas in which improvements can be made to improve the room conditions and indoor air quality in the hospital ward. It may be beneficial for researchers to look further into this result in order to investigate indoor air quality and ward conditions in Hospital Pakar Sultanah Fatimah and for hospital administrators to review these results and take action to address any issues that have been identified.

4.4 Sick Building Syndrome (SBS) Symptom

The data provided in Table 4 shows the experience of a range of symptoms among ward staff in Hospital Pakar Sultanah Fatimah (HPSF) which is potentially a result of exposure to a particular substance, environment, or condition. It is possible that indoor air quality contaminants may have contributed to the sick building symptom (SBS). Poor indoor air quality can be caused by a variety of factors, such as the presence of indoor pollutants,

inadequate ventilation, or high humidity levels, all of which can have negative effects on health and wellbeing.

Table 4: Present Symptom amongst respondents

Present symptom	Sample (s) = 57(%)	
	Yes	No
Fatigue	47(82.5)	10(17.5)
Feeling heavy-headed	35(61.4)	22(38.6)
Headache	42(73.7)	15(26.3)
Nausea, dizziness	15(26.3)	42(73.3)
Difficulties concentrating	23(40.4)	34(59.6)
Itching, burning or irritation of the eyes	19(33.3)	38(66.7)
Irritated, stuffy or runny nose	34(59.7)	23(40.3)
Hoarse, dry throat	30(52.6)	27(47.4)
Cough	33(57.9)	24(42.1)
Dry or flushed facial skin	21(36.8)	36(63.2)
Scaling/ itching scalp or ears	23(40.4)	34(59.6)
Hands dry, itching, red skin	23(40.4)	34(59.6)
Drowsiness	37(64.9)	20(35.1)
Dizziness	32(56.1)	25(43.9)
Irritating of the eye	20(35.1)	37(64.9)
Others	9(15.8)	48(84.2)

Fatigue was reported by over 80% of individuals in the sample. This high percentage suggests that fatigue may be indicative of an underlying health issue or exposure to a harmful substance or environmental factor in the hospital ward. Commonly, fatigue can be caused by a variety of factors, such as lack of sleep, stress, nutritional deficiencies, or underlying health conditions. Poor indoor air quality can lead to fatigue and lethargy. Moving on to the symptom with the median reported percentage, difficulties concentrating were reported by 40.4% of individuals in the group. This percentage suggests that difficulties concentrating may be a moderate concern for the group, but not as common as other symptoms such as fatigue or headaches. Looking at the symptom with the lowest reported percentage, the "Others" category was reported by 15.8% of individuals in the group. While it is difficult to determine exactly what symptoms fall under this category, the low percentage suggests that these symptoms are less common than others reported in the data. However, it is important to note that any unusual or concerning symptoms should be reported to a healthcare professional, regardless of how many other individuals in the group may be experiencing them. The other symptom recorded in this survey are feeling heavy-headed at 61.4%, headache (73.7%), nausea (26.3%), itching, burning or irritation of the eyes (33.3%), irritated, stuffy or runny nose (59.7%), hoarse, dry throat (52.6%), cough (57.9%), dry or flushed facial skin (36.8%), scaling/ itching scalp or ears (40.4%), hands dry, itching, red skin (40.4%), drowsiness (64.9%), dizziness (56.1%) and irritating of the eye with (35.1%). Many of the symptoms reported in the data are consistent with respiratory or allergic reactions, such as irritation of the eyes or nose, dizziness, and

difficulty breathing. These symptoms may be caused by exposure to pollutants or irritants in the air, such as chemicals or allergens.

The relation between SBS symptoms and building condition was conducted via Mann – Whitney U test.

- i. Step 1: state hypothesis
 H_0 : the present symptoms and building condition are equal
 H_1 : the present symptoms and building condition are not equal
- ii. Step 2: level of significance - The significance level is $\alpha = 0.05$ and the test is two-tailed
- iii. Step 3: Identify the test - Mann – Whitney U test

Decision rule: if the calculated U value is less than or equal to the critical U value, then reject H_0 . From the Critical Values of Mann Whitney U Test Table using $\alpha = 0.05$, $n_1 = 16$, $n_2 = 12$, Critical U = 53. Table 5 shows the rank for the present's symptoms and workplace condition.

Table 5: Rank for Symptom and Workplace Condition.

Present (n_1)	Symptom Rank	Workplace condition (n_2)	Rank
47	28	20	12.5
35	25	27	19
4	1.5	42	27
15	4.5	30	20.5
23	17	19	9.5
19	9.5	17	7
34	24	21	14.5
30	20.5	4	1.5
33	23	19	9.5
21	14.5	19	9.5
23	17	15	4.5
23	17	16	6
37	26		
32	22		
20	12.5		
9	3		
Total	265	Total	141

The Mann-Whitney U statistic is then determined:

For present symptoms:

$$U_1 = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R_1$$

$$U_1 = (16)(12) + \frac{16(16 + 1)}{2} - 265 = 63$$

For workplace conditions:

$$U_2 = n_1 n_2 + \frac{n_2(n_2 + 1)}{2} - R_2$$

$$U_2 = (16)(12) + \frac{12(12 + 1)}{2} - 141 = 129$$

Select the smaller of the two: $U_1 = 63$

Since the calculated U (63) is greater than the critical U (53), we do not reject H_0 . Thus, we accept H_0 which states the present symptoms and building conditions are equal. We can conclude that, there is relation between the building condition and present symptom among the ward's staff. To conduct further research, the researcher plans to perform IAQ assessments and physical measurements in the ward. Subsequently, a correlation analysis will be conducted to explore the relationship between SBS symptoms and the physical parameters.

5. Conclusions

In conclusion, this study aimed to investigate the relationship between building conditions, specifically indoor air quality, and Sick Building Syndrome (SBS) symptoms among ward staff at Hospital Pakar Sultanah Fatimah. The results of this study revealed that there is a significant association between poor building conditions and SBS symptoms among ward staff. Specifically, the study found that inadequate ventilation, high levels of indoor air pollutants, and poor temperature and humidity control were significant contributors to SBS symptoms. These findings suggest that building conditions and indoor air quality play a crucial role in the health and well-being of healthcare workers. Therefore, it is important for hospitals and other healthcare facilities to prioritize building maintenance and regular indoor air quality monitoring to ensure the safety and health of their staff and patients. In summary, this study highlights the importance of maintaining good building conditions and indoor air quality in healthcare facilities to mitigate the risk of SBS symptoms among healthcare workers.

Acknowledgement

This research was funded by a grant from Ministry of Higher Education of Malaysia (FRGS Grant K432).

References

- [1] Rahardhiman, A., Yudhastuti, R., & Azizah, R. (2020). Microbiology Indoor Air Quality at Hospital During the Covid19 Pandemic. *JURNAL KESEHATAN LINGKUNGAN*, 12(1si), 89. <https://doi.org/10.20473/jkl.v12i1si.2020.89-92>.
- [2] Fonseca, A., Abreu, I., Guerreiro, M. J., & Barros, N. (2022). Indoor Air Quality in Healthcare Units—A Systematic Literature Review Focusing Recent Research. In *Sustainability (Switzerland)* (Vol. 14, Issue 2). MDPI. <https://doi.org/10.3390/su14020967>
- [3] Dikmen, Ç. B., & Gültekin, A. B. (2010). Sustainable Indoor Air Quality (IAQ) in Hospital Buildings. In *Survival and Sustainability* (pp. 557–565). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-95991-5_50
- [4] Dib, S., Kumar, A., & Rahman, S. U. (2021). Sources of indoor air pollution in hospital environments: A review. *Journal of Environmental Management*, 293, 112854. <https://doi.org/10.1016/j.jenvman.2021.112854>
- [5] Bahloul, M., Chaieb, I., Mnif, S., & Hamden, K. (2021). Indoor air quality in healthcare facilities: sources of pollutants, monitoring, and mitigation strategies. *Environmental Science and Pollution Research*, 28(21), 27114–27130. <https://doi.org/10.1007/s11356-021-13678-1>
- [6] Zhang, Y., Li, M., Zhu, J., & Li, C. (2021). Sources and control of indoor air pollutants in hospital environments: A review. *Indoor and Built Environment*, 30(9), 1477–1496. <https://doi.org/10.1177/1420326X211023787>

- [7] Tran, T. A. H., & Nguyen, D. D. (2021). Sources and characteristics of indoor air pollutants in healthcare facilities: A review. *Journal of Environmental Sciences*, 105, 116-129. <https://doi.org/10.1016/j.jes.2021.01.015>
- [8] International Labour Organization (2000). "Encyclopedia of Occupational Safety and Health". 2nd edition Building Syndrome.
- [9] Norhidayah, A., Lee, C. K., Azhar, M. K., & Nurulwahida, S. (2013). Indoor air quality and sick building syndrome in three selected buildings. *Procedia Engineering*, 53, 93–98. <https://doi.org/10.1016/j.proeng.2013.02.014>.
- [10] Kumar N, Foster A.2007. Respiratory health effects of air pollution in Delhi and its neighboring areas, *India.Environ Monit Assess*. 135:313–325.
- [11] Guttikunda SK, Jawahar P.2012. Application of SIM-air modeling tools to assess air quality in Indian cities. *AtmosEnviron*. 62:551–561. <http://doi:10.1016/j.atmosenv.2012.08.074>.
- [12] Ghosh A, Mukherji A.2014.Air Pollution and Respiratory Ailments among Children in Urban India: exploringCausality. *Econ Dev Cult Change*. 63(1):91–222. <http://doi:10.1086/677754>.
- [13] Tobollik M, Razum O, Wintermeyer D, Plass D.2015. Burden of outdoor air pollution in Kerala, India—a first health risk assessment at state level. *Int J Environ Res Public Health*. 12(9):10602–10619. <http://doi:10.3390/ijerph120910602>.
- [14] Gawande U, Khanvilkar A, Kadam S, Potdar G, Salvitthal H.2016.Effects of ambient air pollution on respiratory health of adults: findings from a cross-sectional study in Chandrapur, Maharashtra, India. *Int J Res Med Sci*. 4(5):1546–1557. <http://doi:10.18203/2320-6012.ijrms20161226>.
- [15] Rajak, R., & Chattopadhyay, A. (2020). Short and Long Term Exposure to Ambient Air Pollution and Impact on Health in India: A Systematic Review. *International Journal of Environmental Health Research*, 30(6), 593–617. <https://doi.org/10.1080/09603123.2019.1612042>.
- [16] Zainal, Z. A., Hashim, Z., Jalaludin, J., Lee, L. F., & Hashim, J. H. (2019). Sick Building Syndrome among Office Workers in relation to Office Environment and Indoor Air Pollutant at an Academic Institution, Malaysia. In *Malaysian Journal of Medicine and Health Sciences* (Vol. 15, Issue 3).
- [17] Syahzanan Alwi, N., Haryani Hassim, M., & Ainun Hamzah, N. (2021). Indoor Air Quality and Sick Building Syndrome among Garment Manufacturing Workers in Kota Bharu, Kelantan. In *Malaysian Journal of Medicine and Health Sciences* (Vol. 17, Issue SUPP8).
- [18] Lawrernce, I. D., Jayabal, S., & Thirumal, P. (2018). Indoor air quality investigations in hospital patient room. *International Journal of Biomedical Engineering and Technology*, 27(1/2), 124. <https://doi.org/10.1504/ijbet.2018.093091>.
- [19] Božić, J., Ilić, P., & Ilić, S. (2019). Indoor air quality in the hospital: The influence of heating, ventilating and conditioning systems. *Brazilian Archives of Biology and Technology*, 62. <https://doi.org/10.1590/1678-4324-2019180295>
- [20] Gola, M., Settimo, G., & Capolongo, S. (2019). Indoor Air Quality in Inpatient Environments: A Systematic Review on Factors that Influence Chemical Pollution in Inpatient Wards. In *Journal of Healthcare Engineering* (Vol. 2019). Hindawi Limited. <https://doi.org/10.1155/2019/8358306>.
- [21] Lenzer, B., Rupprecht, M., Hoffmann, C., Hoffmann, P., & Liebers, U. (2020). Health effects of heating, ventilation and air conditioning on hospital patients: A scoping review. *BMC Public Health*, 20(1). <https://doi.org/10.1186/s12889-020-09358-1>.

- [22] Babaoglu, U. T., Milletli Sezgin, F., & Yag, F. (2020). Sick building symptoms among hospital workers associated with indoor air quality and personal factors. *Indoor and Built Environment*, 29(5), 645–655. <https://doi.org/10.1177/1420326X19855117>.
- [23] Zenaide-Neto, H., & Nascimento, J. S. do. (2020). Air quality and microbiological control in a hospital in Paraíba, Brazil. *International Journal of Advanced Engineering Research and Science*, 7(9), 99–108. <https://doi.org/10.22161/ijaers.79.13>.
- [24] Stojanović Bjelić, L., Ilić, P., & Farooqi, Z. U. R. (2020). Indoor Microbiological air Pollution in the Hospital. *Quality of Life (Banja Luka) - APEIRON*, 18(1–2). <https://doi.org/10.7251/qol2001005s>.
- [25] Busso, I. T., Herrera, F., Tames, M. F., Gasquez, I. G., Camisassa, L. N., & Carreras, H. A. (2020). QuEChER method for air microbiological monitoring in hospital environments. *Journal of Infection in Developing Countries*, 14(1), 66–73. <https://doi.org/10.3855/jidc.11563>.
- [26] Angelova, R. A., Markov, D., Velichkova, R., Stankov, P., & Simova, I. (2021). Exhaled carbon dioxide as a physiological source of deterioration of indoor air quality in non-industrial environments: Influence of air temperature. *Energies*, 14(23). <https://doi.org/10.3390/en14238127>.
- [27] Kapalo, P., Domnița, F., Bacoțiu, C., & Spodyniuk, N. (2018). The Impact of Carbon Dioxide Concentration on the Human Health - Case Study. *Journal of Applied Engineering Sciences*, 8(1), 61–66. <https://doi.org/10.2478/jaes-2018-0008>.
- [28] Jaiswal, A. (2021). Effects of Carbon Monoxide and Cyanide Poisoning On Human Health. *Public Health Open Access*, 5(1). <https://doi.org/10.23880/phoa-16000182>.
- [29] Cincinelli, A., & Martellini, T. (2017). Indoor air quality and health. In *International Journal of Environmental Research and Public Health* (Vol. 14, Issue 11). MDPI. <https://doi.org/10.3390/ijerph14111286>.
- [30] Kim, K. H., Jahan, S. A., & Lee, J. T. (2011). Exposure to formaldehyde and its potential human health Hazards. *Journal of Environmental Science and Health - Part C Environmental Carcinogenesis and Ecotoxicology Reviews*, 29(4), 277–299. <https://doi.org/10.1080/10590501.2011.629972>.
- [31] Qu, M., Lu, J., & He, R. (2017). Formaldehyde from environment. In *Formaldehyde and Cognition* (pp. 1–19). Springer Netherlands. https://doi.org/10.1007/978-94-024-1177-5_1.
- [32] Dan, S., Pant, M., Kaur, T., & Pant, S. (2020). TOXIC EFFECT OF FORMALDEHYDE: A SYSTEMATIC REVIEW Technical SCIENCE ☼ View project The Wound Healing Efficacy of AP-Hydrogel for excision wound View project TOXIC EFFECT OF FORMALDEHYDE: A SYSTEMATIC REVIEW. In *@International Research Journal of Modernization in Engineering* (Vol. 179). www.irjmets.com.
- [33] Kraus, M., & Šenitková, I. J. (2020). Level of Total Volatile Organic Compounds (TVOC) in the context of Indoor Air Quality (IAQ) in Office Buildings. *IOP Conference Series: Materials Science and Engineering*, 728(1). <https://doi.org/10.1088/1757-899X/728/1/012012>.
- [34] Mack, S. M., Madl, A. K., & Pinkerton, K. E. (2020). Respiratory health effects of exposure to ambient particulate matter and bioaerosols. *Comprehensive Physiology*, 10(1), 1–20. <https://doi.org/10.1002/cphy.c180040>.
- [35] Lippmann, M. (1989). Health Effects Of Ozone A Critical Review. *Journal of the Air Pollution Control Association*, 39(5), 672–695. <https://doi.org/10.1080/08940630.1989.10466554>.

- [36] Changkaew, K., Homkham, N., Ketsakorn, A., & Wangwongwatana, S. (2020). Assessment of indoor air quality and particle size distribution of total bacteria and Staphylococcus spp. in an urban hospital in Thailand Thailand Air Pollution Regulatory Development View project. <https://www.researchgate.net/publication/346679887>.
- [37] Lutfi, S. M. (2019). Bacteria Effect on Health and Human-Review. In American *International Journal of Biology and Life Sciences* (Vol. 1, Issue 1). www.acseusa.org/journal/index.php/aijbls.
- [38] Zawawi, E. M. A., Azaiz, A. Z., Kamaruzzaman, S. N., Ishak, N. M., & Yussof, F. N. M. (2019). Indoor Air Quality (IAQ) Performance in Refurbished Projects: A Case Study of Two Private Schools in Selangor. <https://doi.org/10.1051/mateconf/2019>.
- [39] Mui, K. W., Chan, W. Y., Wong, L. T., & Hui, P. S. (2007). Fungi - An indoor air quality assessment parameter for air-conditioned offices. *Building Services Engineering Research and Technology*, 28(3), 265–274. <https://doi.org/10.1177/0143624407081507>.
- [40] Almatawah, Q. A., Al-Rashidi, M. S., Yassin, M. F., & Varghese, J. S. (2022). Microbiological contamination of indoor and outdoor environments in a desert climate. *Environmental Monitoring and Assessment*, 194(5). <https://doi.org/10.1007/s10661-022-10032-9>.
- [41] World Health Organization (2010). WHO Guideline for Indoor Air Quality: Selected Pollutant. Copenhagen: World Health Organization.
- [42] Occupational Safety and Health Association USA (2012). OSHA: Occupational Chemical Database. Retrieved on June 2012 from <http://www.osha.gov/chemicaldata/>.
- [43] National Institute for Occupational Safety and Health. (2007). NIOSH Pocket Guide to Chemical Hazards. Pittsburgh: DHHS (NIOSH) No. 2005-149.
- [44] Department of Occupational Safety and Health. (2010). Industry Code of Practice on Indoor Air Quality. Malaysia: JKKP DP(S) 127/379/4-39.
- [45] Hodgson, M. (2002). “Indoor Environmental Exposure and Symptoms”/ Environmental health perspective 110: 663-667.
- [46] HPSF. (2012). Main Page Hospital Pakar Sultanah Fatimah. (Last Updated 2012). <https://hpsf.moh.gov.my/index.php/en/about-us/introduction.html#background>

REAL TIME PATIENT VITAL MONITORING AND ALARM SYSTEM WITH PREDICTION OF ANOMALIES AND FUTURE CLINICAL EPISODES USING MACHINE LEARNING MODELS

Arko Pal^{1*}
Kunal Mujoo²
Ruby Dinakar³
Khushi Gupta⁴
Akarsh Siwach⁵

^{1,2,3,4,5} Department of Computer Science PES University Bangalore, India

²Email: arko.pal2000@gmail.com

³Email: mujoo.kunal@gmail.com

⁴Email: rubydinakar@pes.edu

⁵Email: akarshsiwach10@gmail.com

Abstract: Heart disease is a leading cause of death in India. Many times, heart diseases do not show symptoms until they become a matter of concern. This silent nature calls for real time monitoring of vitals. In this paper, we propose a healthcare monitoring system which makes use of existing wearable devices to measure vitals of the user's body. The data gathered is used to perform real time analysis to detect any irregularities and simultaneously to predict if there are any underlying problems that might be of future concern.

Keywords: Real Time Vital Monitoring, Wearable Devices, Machine Learning, Disease Prediction

1. Introduction

Heart attack is one of the prime causes of death world- wide and a global health problem. Every year approximately 17.9 million people die of a heart attack. According to the WHO, about 23 million people may suffer from heart attacks until 2030. To reduce a person's risk to heart attacks, we need to be able to detect their susceptibility to it. The growing unhealthy and irregular lifestyle and food habits of people have significantly increased the chances of suffering from a heart disease, thereby putting one's life in danger.

Even though heart attacks are life threatening, it may have early symptoms. Early detection and reporting in a timely manner to the health care facilities about these symptoms could greatly help in saving many lives and avoiding consequences.

Global cardiovascular disease risk assessment estimates the 10-year risk, but it does not capture the dynamic changes in personalised risk that closely follow lifestyle habits. Moreover, the cost associated with it is quite high which proposes the need for a low-cost solution.

The coronavirus disease pandemic (2019) saw a significant increase in the need for remote and decentralised patient care. There was an immediate surge in the need to balance medical emergencies with infectious risk. The alarming rate of increasing cross infections that occurred

Copyright © 2023 Global Academic Excellence (M) Sdn Bhd All Rights Reserved

while travelling to and attending community clinics demanded for a remote real-time monitoring tool. There is evidence to suggest that during pandemic patients with acute cardiovascular diseases did not receive appropriate treatment due to fear of reporting to a hospital and delay in setting up an overwhelmed hospital setup at home.

One of the most fatal types of cardiac disease is myocardial infarction (MI), also known as a heart attack. One of the coronary arteries that supplies the heart muscle with oxygen-rich blood becomes blocked as a result.

According to [3] It affects 805,000 persons in the USA each year. Among them, 200,000 people have repeated heart attacks and 605,000 people have their first heart attack. Recurrent heart attacks have a six-fold greater mortality rate than do initial attacks. Additionally, 20% of heart attacks are silent, which means that the victims are not aware of the harm being done. Additionally, if the sufferers receive therapy within 90 minutes of the MI starting, they have a better probability of surviving.

The requirement for real-time monitoring and MI identification is underscored by the quiet nature of MI, greater risk of recurrent MI, and prompt need for MI treatment. Patients with MI are currently observed and managed in a clinical setting using sophisticated medical technology. Due to the devices' portability issues and high battery consumption, a solution like this is not practical for daily and real-time MI monitoring.

Since they do not have these constraints, wearables are suited for remote and real-time monitoring. Utilizing modern technologies enables you to implement cutting-edge healthcare systems. These wearables have numerous sensors and considerable computational power, facilitating for the monitoring and use of health insights. Wearables haven't yet been used well for remote healthcare, though.

1. Hardware which involves detection using sensors and noise removal from the received signals and communicates it to the decision-making system
2. Software device which makes decision from the signal. This corresponds to the Machine Learning component which involves preprocessing, feature selection, training the models, prediction, and measuring the accuracy of prediction by the models.

Machine learning classification algorithms for heart attack detection requires several steps as shown:

1. Preprocessing- A crucial phase in the design and creation of a reliable real-time patient vital monitoring and prediction system is the preprocessing of data.
2. Feature selection- The process of feature selection involves choosing attributes for the master feature vector that will be used to train the ML classifier. Additionally, feature selection techniques are frequently used to optimise the features extracted and are most pertinent for a real-time patient vital monitoring system. Efficiency is increased, and classification errors are decreased.
3. Classifiers- In the last phase, a machine learning (ML) classifier is trained using the master feature vector to make predictions and determine whether the patient is likely to develop a heart condition, enabling early detection and preventative measures.

In the proposed system we present health monitoring system that makes use of the data generated by the wearable devices. Relevant machine learning techniques are applied on the

Copyright © 2023 Global Academic Excellence (M) Sdn Bhd All Rights Reserved

gathered data to perform real time analysis through big data techniques to detect any irregularities and simultaneously to predict if there are any underlying problems that might be of future concern. When the values generated cross a certain threshold value an alert is sent to the medical authority along with the patient's remote family thereby safeguarding patient health.

2. Literature Review

Healthcare is a critical field that has a need for immense accuracy and quick development. With the recent technological advancement there has been a massive growth in the wearable device domain due to higher availability of IOT enabling components, as well as breakthroughs in wireless communication technology. This makes wearable devices more feasible for real time monitoring of patient vitals.

According to [4], Data Science Technique is used for data collection and processing so that this data can be visualised in real time. They also mention an Improved Pigeon Optimisation Algorithm for data collection, compilation, and storage for enhancing the speed of forecasting. In addition to that they use a BS-DNN to monitor the health of the patient wearing the IoT device. The advantages of real-time perspective are that it offers information on time subject to the present user condition at present. However the paper does not predict susceptibility to underlying diseases and potential risks. This study lacked the implementation of a system that utilises the data collected over a long term for prediction of potential risks.

Research by [1], March 2021 focus on using the IoT sensors and Machine Learning Algorithms and analyses machine learning algorithms for prediction and classification of data collected from wearable devices. The authors compile all the results of the various Machine Learning models used and list their positives and negatives. It is concluded that KNN is the most accurate classification algorithm but may take longer to work in real time applications. So, it states that a collection of LSTM and RNN models may improve the performance in real time.

Another research [5] defines a model for remote monitoring and alarm system using IOT wearable device. It collects biomedical data such as pulse, ECG, PPG, temperature and accelerometer. The term IoMT has been used which stands for Internet of Medical Things which has huge potential in the healthcare industry. It integrates IoMT in a cloud environment along with a Kernel Multiview Canonical Correlation Analysis (KMCCA) using SVM and KNN to perform multi vital correlation and prediction. Intelligent apps can be used to connect users to doctors when either of them are in remote locations.

Based on the pros and cons of related works, we use Preprocessing to reduce size of multivariate data [1]. Alert system to notify about anomalies in vitals. Machine Learning models to predict and classify if the user is susceptible to any underlying condition that may be of future concern [1]. Cloud storage to make data available at any time [4]. The prediction model for efficient forecasting. Real time monitoring to know the user's condition at the present time [2].

3. Proposed Methodology

This section describes in detail the methodology used for Real time patient Vital monitoring System. The architecture of the proposed system where the interaction of the users and the system along with its components are shown in the figure below.

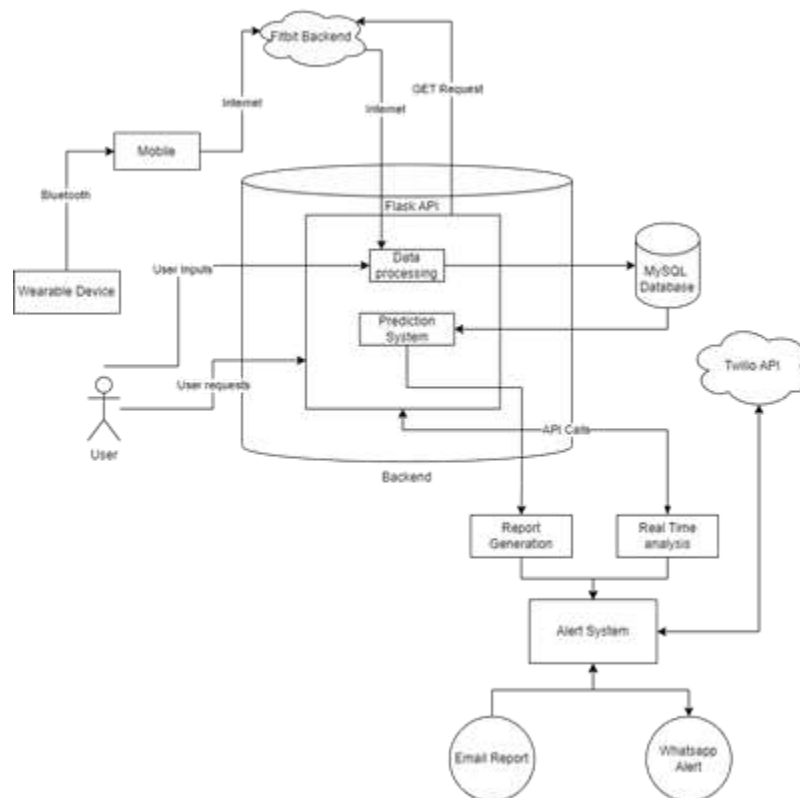


Figure 1. System Architecture

In our application, we have 2 data producers, the User and the Wearable device. The user interacts with the front-end, and all the operations on the frontend are done by the backend. The backend houses an API that responds to all the requests sent by the user. The user enters data such as chest pain, etc from the UI and the remaining data is fetched from the Fitbit API. The user entered data is relevant for 5-6 months since the blood profiling doesn't need to be done that often and hence only needs to be entered 1-2 times a year. The remaining heart data is automatically captured by the wearable and doesn't need any user intervention. Once the data is received, we process the data and store it in the database in a form that is required by our model to predict the output. This data is then used by the prediction system using a Sequential Neural Network to predict if the user is susceptible to a heart disease or not. Once predicted, a report is generated. An external real time analysis system keeps fetching data from the API and based on certain thresholds produces an alert. Both the report and the alert are sent to the user by an Alert system built with the help of Twilio API. The steps involved for Real time patient Vital monitoring System are-

- Model generation using data corpus for training
- Using the model on the data generated from the Fitbit API and the user entered data.
- Alert system to notify about anomalies in vitals

The High level design is given in Figure 2 and further explanation of the steps is given in detail in the subsequent sections.

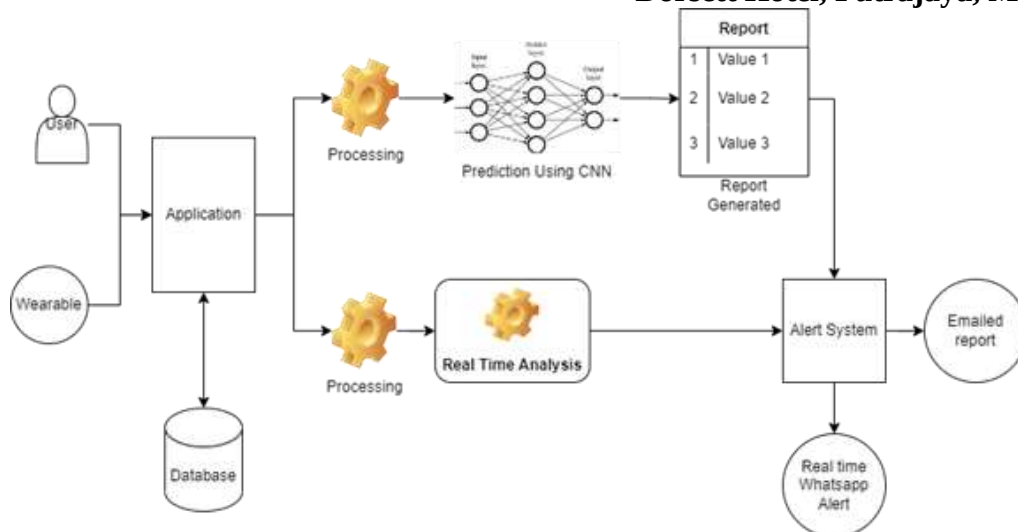


Figure 2. High Level Design of The System

3.1 Model Generation Using Data Corpus for Training

3.1.1 Data Set

The Heart Attack Analysis & Prediction Data set corpus was used for training our model. The corpus comprises of 303 rows, 14 columns. age - Age of the patient(1), sex - Sex of the patient (2), cp - Chest pain type 0= Typical Angina where 1 = Atypical Angina, 2 = Non- anginal Pain, 3 = Asymptomatic (3), trtbps - Resting blood pressure (in mm Hg) (4), chol - Cholesterol in mg/dl (5), fbs - fasting blood sugar greater than 120 mg/dl where 1 = True, 0 = False (6), restecg - Resting electrocardiographic results 0 = Normal, 1 = ST-T wave normality, 2 = Left ventricular hypertrophy (7), thalachh (12) - Maximum heart rate achieved (8), oldpeak - Previous peak (9), slp - Slope (10), caa - Number of major vessels (11), thall - Thallium Stress Test result

Performing the correlation analysis of the attributes to the target output is shown in the figure below:

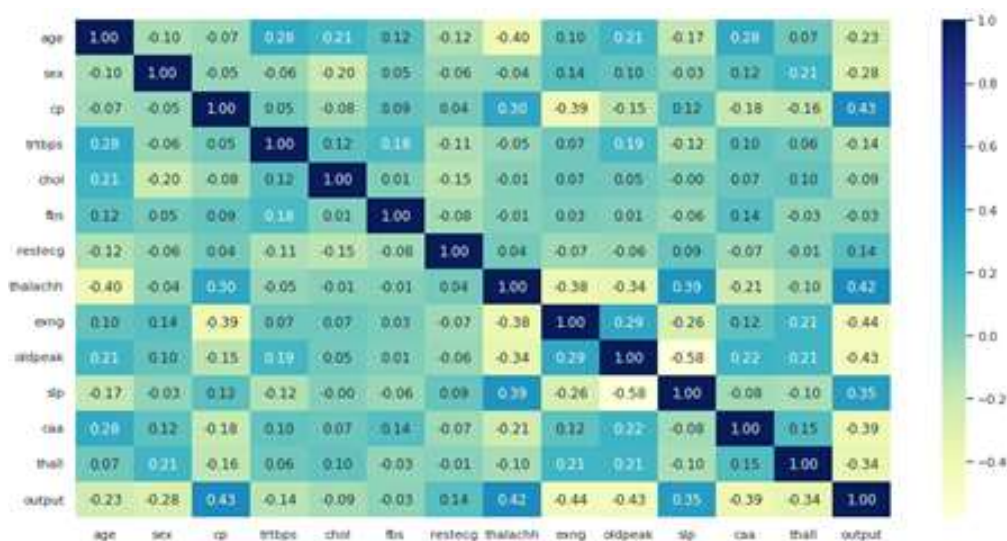


Figure 3. Correlation Heatmap

From the following heatmap we can see that, attributes like exng, oldpeak, caa and thall have very low correlation and barely affect the output. As a result, they have been dropped.

3.1.2 Data Collection and Preprocessing

The data collected is a combination of user inputs and the data collected from the Fitbit API. Once the device is registered to an account in the application, user data and vitals are collected and stored in the database.

Before storing in the database, the data is converted to binary or numerical form required to perform the prediction using a sequential Neural Network. A few insights of the dataset are as follows:

1. There are no NaN values in the data.
2. There are certain outliers in all the continuous features.
3. There are more than 2 times the number of males than females.
4. Figure 3 suggests that there is some correlation in output and cp, and thalachh and slp.
5. Although it makes sense that older persons might have a larger risk of heart attack, Figure 4 shows that this isn't the case.
6. According to Figure 4, higher maximum heart rate may indicate increased susceptibility to heart attacks.
7. Bivariate analysis tells about the following –
 - Non-Anginal chest pain, leads to more odds of having a heart attack.
 - Males have higher chance of heart attack.
 - Thall greater than 1 suggests higher chance of heart attack.
 - People without exercise induced angina, can also be at a risk of having a heart attack.

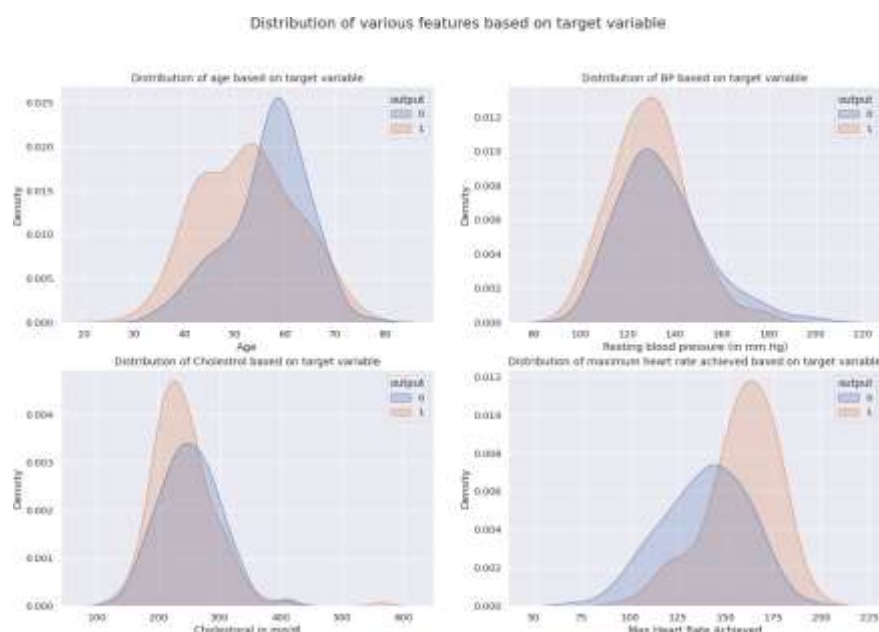


Figure 4. Distribution of Various Features Based on Target Variable

3.1.3 Proposed Model

The dataset was split into train and test set using Stratified Shuffle Split. Several models were generated on the preprocessed dataset, the result obtained is shown below-

	Model	best_score	best_params	type
4	KNN	0.808197	{'algorithm': 'auto', 'n_neighbors': 28, 'p': ...}	oneHotEncoder
0	LogisticRegression	0.788525	{'C': 0.1, 'penalty': 'l2'}	oneHotEncoder
1	SVC	0.786885	{}	oneHotEncoder
9	KNN	0.784426	{'algorithm': 'auto', 'n_neighbors': 29, 'p': ...}	labelEncoder
6	SVC	0.782787	{}	labelEncoder
3	GradientBoostingClassifier	0.782787	{}	oneHotEncoder
5	LogisticRegression	0.782787	{'C': 0.1, 'penalty': 'l2'}	labelEncoder
2	RandomForestClassifier	0.780328	{}	oneHotEncoder
8	GradientBoostingClassifier	0.783934	{}	labelEncoder
7	RandomForestClassifier	0.741803	{}	labelEncoder

Figure 5. Results

Highest accuracy was obtained from KNN model, thus KNN was used as the classification model for our project. ROC curve for the chosen model is shown below.

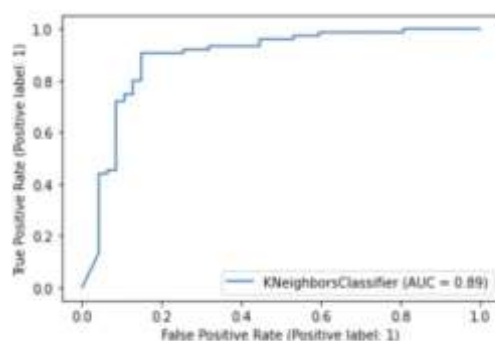


Figure 6. ROC Curve

3.2 Using the Model on The Data Generated from The Fitbit API and the user entered data –

The models are populated from the data obtained from the Fitbit API and the data entered by the user. The user is supposed to register himself and manually enter the values of certain parameters - such as chol, fbs, restecg, thalachh, which remains constant for a long period of time. The remaining parameters which keep changing are obtained every minute from the Fitbit API. The email id of the user is the primary key and has to be unique. Once the output is predicted, a PDF report is created given in Figure 6 is sent to the user.



Health Report

Date: 12/03/2022

Patient Name: Arko Pal
Sex: Male
Age: 31
Email: arko.pal2000@gmail.com
Report Created at: 22:56:35

Test Report

Based on the Inputs received these are your readings:

Blood Cholesterol	140
Fasting Blood Sugar	0
Blood Pressure	115
Maximum Heart Rate Recorded	129
Resting ECG Category	Normal
Chest Pain type	Typical Angina
Slope	1
Prediction	Not susceptible

FINAL REPORT

After Analysing your vitals we predict, that you are not susceptible to a Heart related disease.

Note: This is report not meant for actual diagnosis and is part of the capstone project, we recommend you to visit the doctor in case of any discomfort

Page 1/1

Figure 7. Generated PDF report

3.3 Alert System to Notify About Anomalies in Vitals

The data collected from Fitbit isn't real time therefore a forecasting algorithm had to be used in order to predict the heart rate based on the given trends and hence try to achieve almost real time data to perform analysis. For this we use the forecasting model prophet.

Once the data is collected and based on various studies like [3] & [6] thresholds are set and the data is checked if it is in the normal range. Else a whatsapp alert is created and sent to the phone number of the user/emergency contact.

For testing purposes, Twilio API's whatsapp sandbox feature was used to send the alert to a preset number. For production purposes this alert would be sent to the user's provided number.

4. Conclusion and Future Work

Currently, our system was implemented in a local environment. However, in the future we plan to deploy this system on a cloud platform to make it scalable, and to improve availability, and to make it capable of being used at a national or global scale.

Our model for Real Time monitoring will be of remarkable use in case of any unforeseen pandemic like the Covid 19, in the future. Especially with a country with a population as large as India where the doctor to patient ratio is extremely less. This system will enable doctors to monitor at least 40-50 patients at a time remotely, which is critical to ensure doctor and patient safety.

There is significant scope for improvement in the implementation of our product which could not be implemented due to time constraints but given time and effort this can be a scalable approach to improve doctor-patient safety and scalability in the healthcare industry.

References

- [1] Amani Aldahiri, Bashair Alrashed, and Walayat Hus- sain. Trends in using iot with machine learning in healthprediction system. *Forecasting*, 3(1):181–206, 2021.
- [2] Abdur Rahim Mohammad Forkan and Ibrahim Khalil. Peace-home: Probabilistic estimation of abnormal clini- cal events using vital sign correlations for reliable home-based monitoring. *Pervasive and Mobile Computing*, 38:296–311, 2017.
- [3] Kim Fox, Jeffrey S Borer, A John Camm, Nico-las Danchin, Roberto Ferrari, Jose L Lopez Sendon, Philippe Gabriel Steg, Jean-Claude Tardif, Luigi Tavazzi, Michal Tendera, et al. Resting heart rate in cardiovascular disease. *Journal of the American College of Cardiology*, 50(9):823–830, 2007.
- [4] Rasha M Abd El-Aziz, Rayan Alazani, OsamaR Shahin, Ahmed Elhadad, Amr Abozeid, AhmedI Taloba, and Riyad Alshalabi. An effective data sciencetechnique for iot-assisted healthcare monitoring system with a rapid adoption of cloud computing. *Computa- tional Intelligence and Neuroscience*, 2022, 2022.
- [5] Nonita Sharma, Monika Mangla, Sachi Nandan Mo- hanty, Deepak Gupta, Prayag Tiwari, MohammadShorfuzzaman, and Majdi Rawashdeh. A smartontology-based iot framework for remote patient mon- itoring. *Biomedical Signal Processing and Control*, 68:102717, 2021.
- [6] Dongfeng Zhang, Weijing Wang, and Fang Li. Asso- ciation between resting heart rate and coronary artery disease, stroke, sudden death and noncardiovascular diseases: a meta-analysis. *Cmaj*, 188(15):E384–E392, 2016.

INFLUENCE OF TiO_2 AND Al_2O_3 NANOPARTICLES ADDITION ON THE IMC GROWTH OF SN-3.0AG-0.5CU LEAD-FREE SOLDER AT DIFFERENT REFLOW TEMPERATURE

Nur Haslinda Mohamed Muzni ^{1*}
Ervina Efzan Mhd Noor ²
Mohd Mustafa Al-Bakri Abdullah ³

^{1,2} Faculty of Engineering and Technology, Multimedia University, 75450 Ayer Keroh, Malacca, Malaysia

³ Faculty of Chemical Engineering & Technology, Universiti Malaysia Perlis, 02600 Arau, Perlis, Malaysia

Abstract: This study investigated the influence of adding 0.50 wt% and 1.0 wt% titanium oxide (TiO_2) and aluminium oxide (Al_2O_3) nanoparticles on the properties of Sn3.0Ag0.5Cu (SAC305) lead-free solder alloy. In terms of thermal properties, SAC305 lead-free solder with or without addition of TiO_2 and Al_2O_3 nanoparticles, the melting temperature is very similar and comparable. The wetting performance of the SAC305 nanocomposite solder was determined by the contact angle. The best contact angle recorded is 31.65° at reflow temperature of 250°C for 0.50 wt% addition of TiO_2 and Al_2O_3 nanoparticles. The thickness of the IMC layer was observed using an Optical Microscope and measured. Results show that the thickness of the IMC layer decreased when TiO_2 and Al_2O_3 nanoparticles were reinforced into the SAC305 solder system, which shows significant decrease when 0.50wt% of TiO_2 and Al_2O_3 nanoparticles were added. However, beyond 0.50 wt% the thickness of IMC layer increases along with increasing reflow temperature by 25%. The results reveal that the adsorption effect of TiO_2 and Al_2O_3 nanoparticles will restrict the growth of the IMC layer under controlled concentration of the nanoparticles and reflow temperature. The optimal parameters to enhance the performance of SAC305 solder by reinforcing TiO_2 and Al_2O_3 nanoparticles are reflowed at 250°C at 0.50 wt% concentration of TiO_2 and Al_2O_3 nanoparticles.

Keywords: Lead-Free Solder, Nanoparticles, Intermetallic Compound, Nanocomposite

1. Introduction

Tin-lead (Sn-Pb) solders are being widely used due to they are cost-effective [1], possess low melting point [2], good wettability [3,4] and have distinctive mechanical properties [5,6] in most of all in the electronic packaging industries [1,4]. Owing to environmental consequences (toxicity of Pb), thus, it contributed to a ban on electronics products in majority of countries. The use of harmful substances in electronic products have been prohibited by the European Union's Restriction of Hazardous Substances (RoHS) in 2006 [1,7]. As a result of the lead restriction in the electronic industry, there has been an increasing number of lead-free solders been investigated and developed to replace them [8,9]. Sn-Ag-Cu lead-free solder is one of the most promising lead-free solders due to its superior reliability, creep resistance and thermal fatigue properties [10].

Lead-free candidate solders based on Sn-Ag-Cu alloys are promising for variety of applications [11,12]. The properties of interfacial IMC layers are constantly being explored and improved

in order to improve the reliability of Sn-Ag-Cu solder joints. Solders with minor alloying is one other method for improving the performance of the interfacial IMC layers in solder system. Various IMCs are formed from all common substrate materials in electronic products during soldering [11]. Furthermore, understanding on the solder/substrate interface evolution process and mechanism in the solder joints is also critical for understanding reliability issues from the metallurgical standpoint and optimizing the soldering process.

The use of ceramics as reinforcement in the solder joints have been a novel trend in the development of solder alloys. The use of ceramic reinforcement in nanoparticle form usually below 100nm can promote heterogeneous nucleation. A wide selection of ceramics nanoparticles such as TiO_2 , ZnO , ZrO_2 , Al_2O_3 , La_2O_3 , SiC and Fe_2O_3 can be used as reinforcement to create lead-free nanocomposite solder alloy [13,14]. As a result, there has been an increased demand for the development of high-performance lead-free solder that can strengthen the characteristics they possess. Multiple research efforts have revealed that one of the potential solutions is producing nanocomposite solders by incorporating a relatively small number of nanoparticles into SAC solder [12].

2. Methodology

SAC305 solder was produced with 96.5% Tin (Sn), 3.0% Argentum (Ag), and 0.5% Copper (Cu) in weight percentages (wt%). The SAC305 lead-free nanocomposite solder was prepared by adding 0.50 wt% and 1.0 wt% of TiO_2 and Al_2O_3 nanoparticles respectively into the lead-free SAC305 nanocomposite solder at 250°C in a graphite crucible for 1 hour, and these nanocomposite solders were designated as Sn-3.0Ag-0.5Cu-0.50 TiO_2 -0.50 Al_2O_3 and Sn-3.0Ag-0.5Cu-1.0 TiO_2 -1.0 Al_2O_3 . Then, the prepared solder alloy was cut into 5mm diameter billet with 1mm thickness. The wettability of the nanocomposite solders was evaluated by measuring their contact angles. The nanocomposite solders were melted on copper (Cu) substrate at size of 10x10mm at different reflow temperatures of 230, 250, 270 and 290°C on a laboratory hotplate for 60s. The contact angle was examined by using optical microscope and measured using VIS Pro software. The thickness of the intermetallic compound layer was obtained using an optical microscope. The melting details of the lead-free SAC305 nanocomposite solders were investigated using Differential Scanning Calorimetry (DSC, Perkin Elmer). In an Argon atmosphere, each solder alloy was heated at a rate of 10°C/min to temperatures ranging from 30-300 °C.

3. Results

3.1 Thermal Properties

Figure 1 shows the differential scanning calorimetry (DSC) curves for specimens reinforced with 0.50 and 1.0 wt% composition of TiO_2 and Al_2O_3 nanoparticles are shown in Figure 1. Figure 1(a) shows the DSC curve of SAC305 lead-free solder without any addition of TiO_2 and Al_2O_3 nanoparticles. Figure 1(b) and (c) shows the DSC curves for SAC305 lead-free solder reinforced with 0.5 and 1.0 wt% of TiO_2 and Al_2O_3 nanoparticles respectively. The melting temperature of SAC305 lead-free solder was measured to be 222.5 °C during the heating process of DSC analysis. The melting temperature was slightly changed 222.6 to 231.2 °C for SAC305 nanocomposite solder as shown in Figure 1(b) and (c). The slight increase in melting temperature was due to high melting point of TiO_2 (TM=1843°C) and Al_2O_3 (TM=2072°C) nanoparticles. Since both TiO_2 and Al_2O_3 nanoparticles has high melting point, it suggests that this is due to the local dissolution of TiO_2 and Al_2O_3 nanoparticles [15]. SAC305 solder and SAC305 nanocomposite solders begin to melt at 217°C, reaching its fully liquid state in range between 228.5 to 233.5°C. It can be observed that the solidus (TS) for all SAC nanocomposite

solders, the solidus temperature was similar to the lead-free SAC305 solder at 217 °C. Meanwhile, the liquidus (TL) temperature of SAC305 nanocomposite solder are increasing slightly from 231 to 233.5 °C as shown in Figure 1(b) and (c). Similar findings were also found that the melting point of SAC lead-free solder increase slightly when reinforced with TiO₂ and Al₂O₃ nanoparticles [16,17]. Yet, the slight increment in the melting temperature is not a major concern as it still falls under the typical reflow temperature range for Pb-free solder which is 240-250°C.

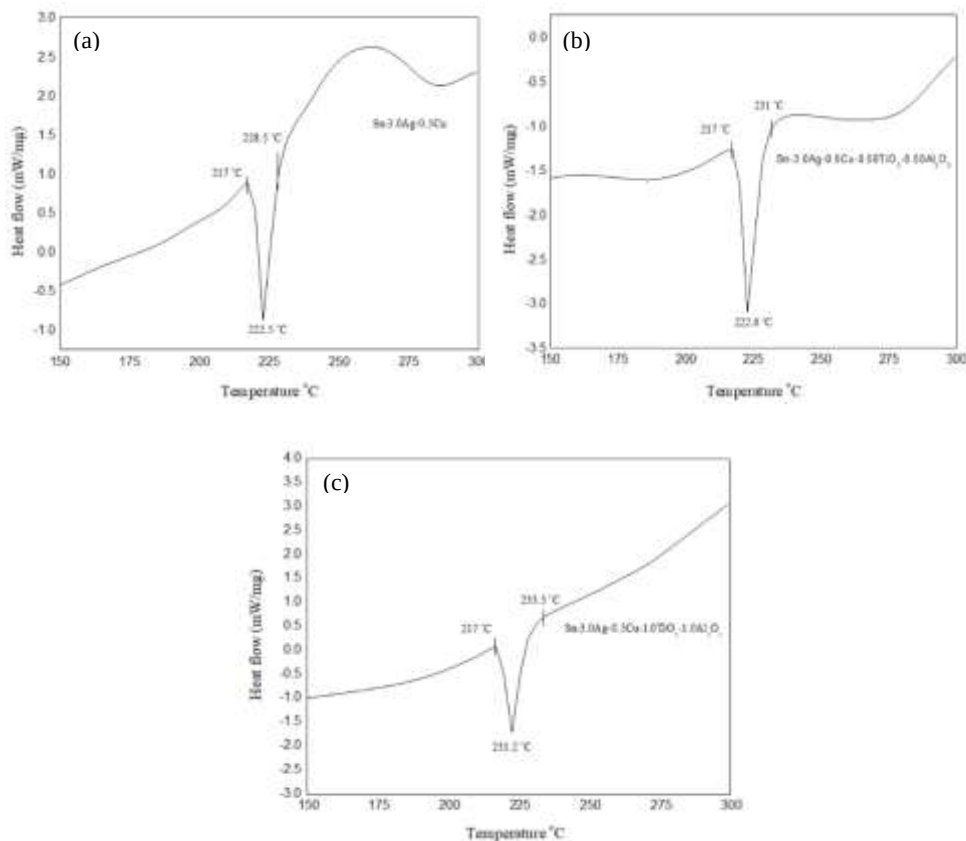


Figure 1: DSC Curves of SAC305 Solder with Different wt% of TiO₂ and Al₂O₃ Nanoparticles Concentrations, (A) SAC305, (B) Sn-3.0Ag-0.5Cu-0.5TiO₂-0.5Al₂O₃ and (C) Sn-3.0Ag-0.5Cu-1.0TiO₂-1.0Al₂O₃

3.2 Wettability

The wettability of the solder, or its capacity to spread molten solder onto a substrate and maintain contact with it, can have an impact on the solder joint's reliability. The contact angle between the molten solder and the Cu substrate can be used to determine the wettability of the solder. Solder has very good wettability when the contact angle is between 0° and 20°. When the value of the wetting angle is between 20° and 40°, it is deemed acceptable and good [10]. Temperature plays a crucial role in determining wettability, spreading characteristics, and interfacial morphology. A good wetting behaviour of the solder by means the contact angle measurement can be achieved by controlling the temperature. Table 1 and Figure 2 present the SAC305 nanocomposite solder's wettability behaviour under a variety of contact angles (Θ) and reflow temperatures, respectively. As shown in Figure 2, the best contact angle recorded is 31.65° at reflow temperature of 250°C for 0.50 wt% addition of TiO₂ and Al₂O₃ nanoparticles. The largest contact angle 46.09° was observed at 290°C for SAC305 solder. Decrease in contact angle from 230°C to 250°C was initially observed, yet an increasing trend was noticed afterward when the reflow temperature increase from 250°C to 290°C. According

Copyright © 2023 Global Academic Excellence (M) Sdn Bhd All Rights Reserved

to these findings, the nanocomposite solder's wetting performance was enhanced by adding nanoparticles of TiO_2 and Al_2O_3 at concentrations of up to 0.50 weight percent at 250 °C.

In other words, the wetting performance can be increased by a small amount at a limited temperature by adding more nanoparticles. In the presence of the TiO_2 and Al_2O_3 nanoparticles, the decreased surface tension between the molten solder and the Cu substrate may be the cause of this occurrence [18,19]. However, upsurge in the contact angle happen as a result due to higher concentrations of TiO_2 and Al_2O_3 nanoparticles that raised the viscosity of the molten solder. Thus, this finding suggested that 0.50 wt% was the optimal TiO_2 and Al_2O_3 nanoparticles concentration at 250°C for the best result. The relationship between concentration of TiO_2 and Al_2O_3 nanoparticles and reflow temperature on the contact angle was excessive concentration of TiO_2 and Al_2O_3 nanoparticles and higher reflow temperature will increase the surface tension and also the viscosity of the molten solder and thus affecting the wetting performance of the nanocomposite solder.

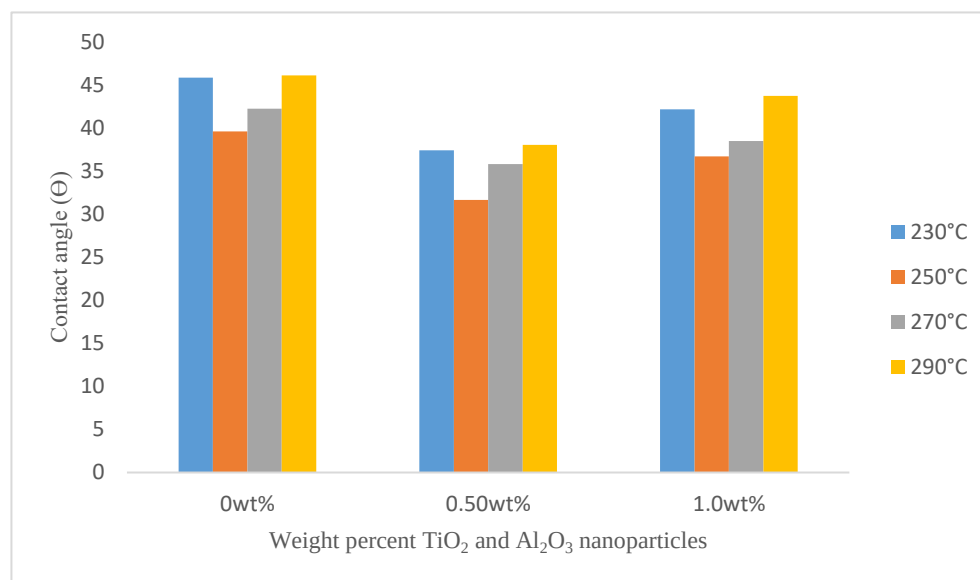


Figure 2: Contact angle of SAC305 nanocomposite solder at different reflow temperature

Table 1: Contact angle of SAC305 nanocomposite solders at different reflow temperature

Sample (°C)	Reflow temperature	Contact angle (Θ)			
		230	250	270	290
SAC305		45.82	45.45	45.97	46.09
Sn-3.0Ag-0.5Cu-0.5TiO ₂ - 0.5Al ₂ O ₃		37.41	31.65	35.81	38.02
Sn-3.0Ag-0.5Cu-1.0TiO ₂ - 1.0Al ₂ O ₃		42.17	36.69	38.49	43.72

3.3 IMC Growth Formation

Figure 3 shows the graph of IMC thickness under different reflow temperature at different concentration of TiO_2 and Al_2O_3 nanoparticles. IMC thickness of the solder was found to decrease at 0.50 wt% under 250°C reflow temperature then increase with increasing reflow temperature up to 290°C. For all SAC305 nanocomposite solder, it was observed that the IMC thickness in the solder is thickest at the highest temperature (290°C) as listed in Table 2. This observation shows that the TiO_2 and Al_2O_3 nanoparticles reacts with Cu substrate as the reflow temperature increases. It can be seen that thinnest IMC layer formed in SAC305-0.50 TiO_2 -0.50 Al_2O_3 solder, a decrease from 1.44 to 1.40 μm at 250°C. When the weight percent of TiO_2 and Al_2O_3 increases, the IMC thickness increase from 1.80 to 1.95 μm for SAC305 solder, 1.44 to 1.55 μm for 0.50 wt% and 3.24 to 3.42 μm for 1.0 wt%. The findings show that the growth of IMC in liquid-state reactions can be inhibited by the addition of TiO_2 and Al_2O_3 nanoparticles, and that this inhibition is most effective when the concentration of TiO_2 and Al_2O_3 nanoparticles is increased to about 0.50 wt% at 250°C. Ceramic nanoparticles give significant enhancements on the properties of lead-free solder alloy include microstructures, thermal behaviours, mechanical properties, and growth of IMCs [10,13,20]. The refined Ag_3Sn IMCs phases proves TiO_2 and Al_2O_3 nanoparticles significantly restrict the growth of IMCs by acting as barrier between solder and substrate [13,14]. Thickness of the IMC layer decreases significantly when reinforced with 0.50 wt% of TiO_2 and Al_2O_3 nanoparticles but further addition of TiO_2 and Al_2O_3 nanoparticles could promotes growth of the IMC [13,14,21]. The thickness of the overall IMC layer decreases with an increase in TiO_2 and Al_2O_3 concentration along with the rise in reflow temperature. The decline trend starts when the weight percentage of TiO_2 and Al_2O_3 is at 0.50 wt%. IMC layer growth is elevated by an incremental increase in TiO_2 and Al_2O_3 concentration up to 1.0 wt%, though it remains thinner than that of the SAC305 solder joints.

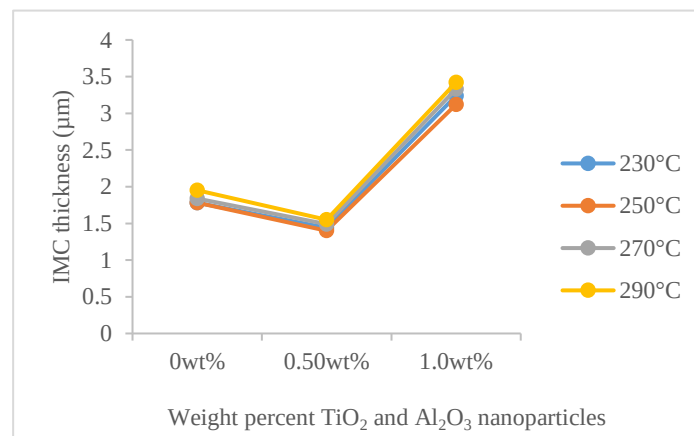


Figure 3: IMC Thickness of SAC305 Nanocomposite Solder at Different Reflow Temperature

Table 2: IMC Thickness of SAC305 Nanocomposite Solder at Different Reflow Temperature

Sample	Reflow temperature (°C)	IMCs thickness (µm)			
		230	250	270	290
SAC305		1.80	1.78	1.84	1.95
Sn-3.0Ag-0.5Cu-0.5TiO ₂ -0.5Al ₂ O ₃		1.44	1.40	1.49	1.55
Sn-3.0Ag-0.5Cu-1.0TiO ₂ -1.0Al ₂ O ₃		3.24	3.12	3.33	3.42

Similar findings by Tsao et al. [14] and Chuang et al. [22] imply that the reinforcement effect of Al₂O₃ nanoparticles is reduced once the amount reaches 1.0 wt%. The results discover that reinforcing TiO₂ and Al₂O₃ nanoparticles will restrict the growth of IMC layer under controlled concentration of TiO₂ and Al₂O₃ nanoparticles and reflow temperature. The growth of the IMC layer is restricted by the increased adsorption of Al₂O₃ nanoparticles, which decreases the surface free energy of the Ag₃Sn IMC grains, in accordance with the theory of the adsorption of surface-active materials [13]. Same goes for TiO₂ nanoparticles where it acted as reinforcements and barrier for solder dislocation [13]. TiO₂ nanoparticles' adsorption effect inhibits the growth of the IMC layer [8,13,23]. This result was in accordance with other findings where the IMC layer thickness decreased by approximately by 50% when reinforcing 0.50 wt% of TiO₂ nanoparticles to SAC305 lead-free solder [13,15]. The adsorption effect of TiO₂ and Al₂O₃ nanoparticles not only restrict the growth of the IMC layer but also serves as a diffusion barrier at the interface between molten solder and Cu substrate.

Conclusions

SAC305-0.50TiO₂-0.50Al₂O₃ and SAC305-1.0TiO₂-1.0Al₂O₃ lead-free nanocomposite solders were successfully produced by reinforcing 0.50 wt% and 1.0 wt% of TiO₂ and Al₂O₃ nanoparticles respectively. At 217°C, SAC305 nanocomposite solders had a solidus temperature that was nearly identical to lead-free SAC solders. The SAC305 nanocomposite solders' solidus and liquidus temperatures rise as the percentage of TiO₂ and Al₂O₃ nanoparticles increases, but still falls under the typical reflow temperature range for Pb-free solder, which is 240–250°C. The wetting performance of the solder system were improved under 250°C at 0.50 wt% concentration of TiO₂ and Al₂O₃ nanoparticles and falls under good standard of wetting angle between 20° < θ < 40°. This occurrence could be because of the reduced surface tension between the molten solder and Cu substrate in the presence of the TiO₂ and Al₂O₃ nanoparticles. On top of that, the IMC layer thickness was also improved under 250°C at 0.50 wt% concentration of TiO₂ and Al₂O₃ nanoparticles. The growth of the IMC layer thickness was restricted and controlled due to the adsorption effects of TiO₂ and Al₂O₃ nanoparticles and their roles as diffusion barrier and reduce the diffusion rate of Cu atoms to dissipate to the molten solder and react with Sn resulting in formation of Cu₆Sn₅ and Cu₃Sn IMC layer. Thus, this finding proposed that the addition of TiO₂ and Al₂O₃ nanoparticles improves the performance of SAC305 solder alloy under controlled concentrations of TiO₂ and Al₂O₃ nanoparticles and reflow temperature.

Acknowledgement

Sincere gratitude to MOHE for funding assistance provided by FRGS Research Grant No. FRGS/1/2018/TKMUE180043.

References

- [1] M.A., Fazal, N.K., Liyana, A., Anas. "A critical review on performance, microstructure and corrosion resistance of Pb-free solders". *Measurement* 134, (2019): 897-905. <https://doi.org/10.1016/j.measurement.2018.12.051>
- [2] Pal, Manoj Kumar, Gergely, Gréta, Koncz-Horváth, Daniel, Gácsi, Zoltán. "Investigation of microstructure and wetting behavior of Sn-3.0Ag-0.5Cu (SAC305) lead-free solder with additions of 1.0 wt % SiC on copper substrate". *Intermetallics* 128, (2021): 106991. <https://doi.org/10.1016/j.intermet.2020.106991>
- [3] M.N., Ervina Efzan, I., Siti Norfarhani. "Effect of aging times on Sn-Ag-Cu solder alloy". *Transactios on Electrical and Electronic Materials* 16, no. 3 (2015): 112-116. <http://dx.doi.org/10.4313/TEEM.2015.16.3.112>
- [4] Pal, Manoj Kumar, Gergely, Gréta, Koncz-Horváth, Daniel, Gácsi, Zoltán. "Characterization of the interface between ceramics reinforcement and lead-free solder matrix". *Surfaces and Interfaces* 20, (2020): 100576. <https://doi.org/10.1016/j.surfin.2020.100576>
- [5] Tikale, Sanjay, Prabhu, K. Narayan. "Effect of Multiple Reflow Cycles and Al₂O₃ Nanoparticles Reinforcement on Performance of SAC305 Lead-Free Solder Alloy". *Journal of Materials Engineering and Performance* 27, (2018): 3102-3111. <https://doi.org/10.1007/s11665-018-3390-y>
- [6] Qiu, Hongyu, Hu, Xiaowu, Li, Shuang, Wan, Yongqiang, Li, Qinglin. "Shear strength and fracture surface analysis of lead-free solder joints with high fraction of IMCs". *Vacuum* 180, (2020): 109611. <https://doi.org/10.1016/j.vacuum.2020.109611>
- [7] Wu, Jie, Xue, Song-bai, Wang, Jing-wen, Liu, Shuang, Han, Yi-long, Wang, Liu-jue. "Recent progress of Sn-Ag-Cu lead-free solders bearing alloy elements and nanoparticles in electronic packaging". *Journal of Materials Science: Materials in Electronics* 27, (2016): 12729-12763. <https://doi.org/10.1007/s10854-016-5407-3>
- [8] Xiong, Ming-yue, Zhang, Liang. "Interface reaction and intermetallic compound growth behavior of Sn-Ag-Cu lead-free solder joints on different substrates in electronic packaging". *Journal of Materials Science* 54, (2019): 1741-1768. <https://doi.org/10.1007/s10853-018-2907-y>
- [9] Zhang, Liang, Cui, Jun-hua, Han, Ji-guang, He, Cheng-wen, Guo, Yong-huan, Yuan, Jian-min. "Finite Element Analysis of SnAgCu(Zn, Co, Fe) Lead-free Solder Joints for Electronic Packaging". *International Journal of Nonlinear Sciences and Numerical Simulation* 15, no. 3-4 (2014): 197-206. <https://doi.org/10.1515/ijnsns-2012-0063>
- [10] Sun, Lei, Zhang, Liang. "Properties and Microstructures of Sn-Ag-Cu-X Lead-Free Solder Joints in Electronic Packaging". *Advances in Materials Science and Engineering* 2015, (2015): 1-16. <https://doi.org/10.1155/2015/639028>
- [11] Zeng, Guang, Xue, Songbai, Zhang, Liang, Gao, Lli, Dai, Wei, Luo, Jiadong. "A review on the interfacial intermetallic compounds between Sn-Ag-Cu based solders and substrates". *Journal of Materials Science: Materials in Electronics* 21, (2010): 421-440. <https://doi.org/10.1007/s10854-010-0086-y>
- [12] Cheng, Shunfeng, Huang, Chien-Ming, Pecht, Michael. "A review of lead-free solders for electronics applications". *Microelectronics Reliability* 75, (2017): 77-95. <https://doi.org/10.1016/j.microrel.2017.06.016>

- [13] Tan, Ai Ting, Tan, Ai Wen, Yusof, Farazila. "Influence of nanoparticle addition on the formation and growth of intermetallic compounds (IMCs) in Cu/Sn–Ag–Cu/Cu solder joint during different thermal conditions". *Science and Technology of Advanced Materials* 16, no. 3 (2015): 033505. <https://doi.org/10.1088/1468-6996/16/3/033505>
- [14] Tsao, L.C., Chang, S.Y., Lee, C.I., Sun, W.H., Huang, C.H. "Effects of nano- Al_2O_3 additions on microstructure development and hardness of Sn3.5Ag0.5Cu solder". *Materials & Design* 31, no. 10 (2010): 4831-4835. <https://doi.org/10.1016/j.matdes.2010.04.033>
- [15] Chang, S.Y., Jain, C.C., Chuang, T.H., Feng, L.P., Tsao, L.C. "Effect of addition of TiO_2 nanoparticles on the microstructure, microhardness and interfacial reactions of Sn3.5AgXCu solder". *Materials & Design* 32, no. 10 (2011): 4720-4727. <https://doi.org/10.1016/j.matdes.2011.06.044>
- [16] Tsao, L.C., Chang, S.Y. "Effects of Nano- TiO_2 additions on thermal analysis, microstructure and tensile properties of Sn3.5Ag0.25Cu solder". *Materials & Design* 31, no. 2 (2010): 990-993. <https://doi.org/10.1016/j.matdes.2009.08.008>
- [17] Alsorory, Hamed, Gumaan, Mohammed, Shalaby, Rizk Mostafa. "Effect of Al_2O_3 Nanoparticle Addition on the Microstructure, Mechanical, Thermal, and Electrical Properties of Melt-Spun SAC355 Lead-Free Solder for Electronic Packaging". *Journal of Materials Engineering and Performance*, (2022). <http://dx.doi.org/10.1007/s11665-022-07752-x>
- [18] Mhd Noor, Ervina Efzan, Mhd Nasir, Nur Faziera, Idris, Siti Rabiatal Aisya. "A review: lead free solder and its wettability properties". *Soldering & Surface Mount Technology* 28, no. 3 (2016): 125-132. <https://doi.org/10.1108/SSMT-08-2015-0022>
- [19] Tikale, Sanjay, Prabhu, K. Narayan. "Performance and reliability of Al_2O_3 nanoparticles doped multicomponent Sn-3.0Ag-0.5Cu-Ni-Ge solder alloy". *Microelectronics Reliability* 113, (2020): 113933. <https://doi.org/10.1016/j.microrel.2020.113933>
- [20] Zhang, Peng, Xue, Songbai, Wang, Jianhao, Xue, Peng, Zhong, Sujuan, Long, Weimin. "Effect of Nanoparticles Addition on the Microstructure and Properties of Lead-Free Solders: A Review". *Applied Sciences* 9, (2019): 2044. <https://doi.org/10.3390/app9102044>
- [21] Tang, Y., Li, G.Y., Pan, Y.C. "Influence of TiO_2 nanoparticles on IMC growth in Sn–3.0Ag–0.5Cu–x TiO_2 solder joints in reflow process". *Journal of Alloys and Compounds* 554, (2013): 195-203. <https://doi.org/10.1016/j.jallcom.2012.12.019>
- [22] Chuang, T.H., Wu, M.W., Chang, S.Y., Ping, S.F., Tsao, L.C. "Strengthening mechanism of nano- Al_2O_3 particles reinforced Sn3.5Ag0.5Cu lead-free solder". *Journal of Materials Science: Materials in Electronics* 22, (2011): 1021-1027. <https://doi.org/10.1007/s10854-010-0253-1>
- [23] Wu, Jie, Xue, Songbai, Wang, Jingwen, Wu, Mingfang, Wang, Jianhao. "Effects of α - Al_2O_3 nanoparticles-doped on microstructure and properties of Sn–0.3Ag–0.7Cu low-Ag solder". *Journal of Materials Science: Materials in Electronics* 29, (2018): 7372-7387. <https://doi.org/10.1007/s10854-018-8727-7>

COMPARATIVE INVESTIGATION OF RESISTANCE PREDICTION FOR SURFACE COMBATANT SHIP MODEL USING CFD MODELING

Alaaeldeen M. Elhadad¹

¹Shipbuilding Engineering Department, Military Technical College, Cairo, Egypt
dr.aladdinahmed@gmail.com

Abstract: The Multi-Mission Surface Combatant (MMSC) is a highly maneuverable combatant ship capable of littoral and open ocean operations. It was designed to confront modern maritime and economic security threats. The MMSC takes the proven capabilities of the littoral combat ship and the inherent flexibility of the Freedom-variant hull to meet the unique maritime requirements of international navies. Computational fluid dynamics (CFD) is applied to present a method to predict the resistance for new surface combatant ship. First, calculations for a typical benchmark DTMB 5415 model are carried out using three different mesh sizes for Froude numbers from 0.10 to 0.45 for the purpose of model validation by StarCCM+. The numerical results are compared with the experimental data and the published CFD solutions in terms of wave field and resistance coefficients for accuracy of the solution parameters. Finally, the method is used to study the influence of Froude number variation on the total resistance and wave pattern for the new combatant ship model under the same conditions. Quantitative agreement between the numerical simulations has been observed. This demonstrates that our CFD model is capable of simulating the steady flow around a ship hull with an acceptable accuracy and thus can be used as a complementary tool to laboratory model tests for ship design and ship hydrodynamic research.

Keywords: Combatant ship; MMSC; DTMB; Resistance; CFD

Introduction

Surface combatant are a significant sector in naval vessels that are built for surface warfare with their own weapons and armed forces. They are large, heavily armed surface ships, including various types of battleships, battlecruisers, cruisers, destroyers, frigates, and corvettes. They are primarily intended to engage with surface, and submerged targets by deploying weapons from the ship rather than by manned carried craft [1]. They can carry out other missions, including, counter-narcotics operations and maritime interdiction. The development of new military abilities to confront potential threats or upgrade the existing capabilities urged the design and construction of new naval warships.

This study aims to introduce a flexible and applicable method to utilize CFD calculations to predict model resistance in the early design stage. CFD was applied to high performance hull form types to determine local flow details, optimize a given design, and select the most promising candidates for further testing [2].

Firstly, the free surface flow around the DTMB #5415 model was calculated on various Froude number cases and verified with the experimental data of the model. Furthermore, a numerical

estimate of the effect of Froude number on wave pattern and total resistance for the new ship model was calculated and compared with confirmed validated data of the model [3].

In this paper, we introduce a method to estimate the hull model resistance and powering, as well as the flow field surrounding a new combatant hull model, by using the StarCCM+ as a primary step for production of a validated ship model in order to carry out additional experimental investigation in a towing tank [4] [5].

Mathematical Model and Governing Equations

The incompressible, viscous turbulent flow field and two-phases (air and water) are governed by the RANS equations:

$$\frac{\partial U_i}{\partial x_i} = 0 \quad (1)$$

$$\frac{\partial u_i}{\partial t} + u_i \frac{\partial u_i}{\partial x_j} = -\frac{1}{\rho} \left(\frac{\partial p}{\partial x_i} \right) + \nu \frac{\partial^2 u_i}{\partial x_j^2} - \frac{\partial \overline{u_i' u_j'}}{\partial x_j} \quad (2)$$

where $i, j = 1, 2$ and x_1, x_2 denote the horizontal and the vertical dimensions, respectively, u_1 and u_2 are the corresponding to the mean velocity components, $\overline{u_i' u_j'}$ defines the Reynolds stress component with u_i' being the fluctuating part of the velocity, p represents the dynamic pressure, and ρ is the fluid density.

SST k- ϵ model was applied to compute the turbulent viscosity meanwhile the multiphase free-surface flow is simulated by the VOF method. By using the Boussinesq approximation, the Reynolds stress component was defined in terms of a turbulent viscosity, ν_T , and the mean flow gradients as follows:

$$-\overline{u_i' u_j'} = \nu_T \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) - \frac{2}{3} k \delta_{ij} \quad (3)$$

where δ_{ij} is Kronecker delta function, and k is turbulent kinetic energy.

In the current study, a three-dimensional numerical simulation was conducted for the flow past the hull model [6]. The numerical simulations were predicted by a two-layer form, developed by StarCCM+ software. The fluid domain is partitioned into a finite number of cells, and the governing equations for fluid flow are then converted by a discretization process into algebraic form to be solved. Using a straightforward technique, the pressure and velocity fields were coupled [7] [8].

Grid Generation, Solution Domain and Boundary Conditions

In this study, choosing a realistic hull model is a crucial step. A very well-known test model, surface combatant DTMB model #5415-51, is mathematically specified with an analytical description. This combatant hull is typically seen as a preferred test sample to verify a new numerical approach. The combatant ship model's hull shape, grid structure, properties, and profile, which correspond to a scale of 24.824, are listed in Figure 1 and Table 1 [9].

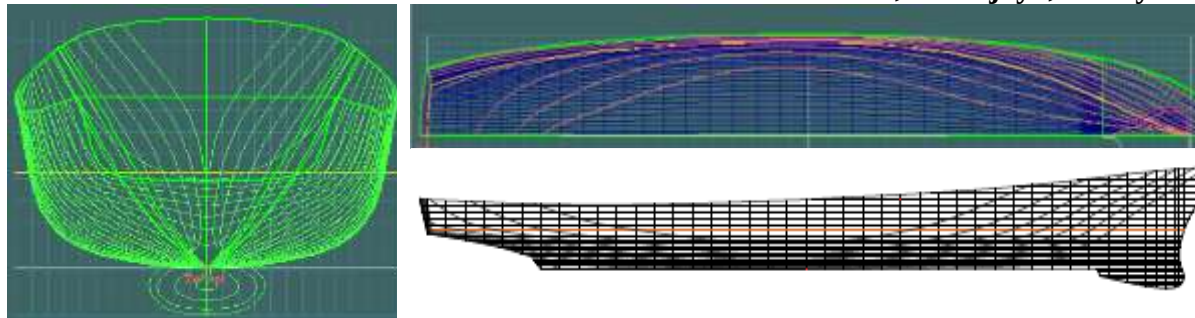


Figure 1(A) Lines Plan (B) Numerical Grid of DTMB #5415-51 Hull Surface

Table 2 Main Particulars of DTMB #5415 (ITTC, 2005)

L_{pp} (m)	5.72	LCG (m)	2.884
B (m)	0.760	VCG (m)	0.056
T (m)	0.248	Wetted Surface Area (S) (m^2)	4.786
Δ (tonne)	0.549	Block Coefficient (C_B)	0.506

The hull mesh domain is separated into the water and air zones with cubical form. The x-axis is set up to point toward the bow, the y-axis to portside, and the z-axis is configured to point upward in the Cartesian coordinate system. The downstream and upstream limits are situated three and one hull length from the stern and bow, respectively. The air zone is half a model length above the water's surface, while the water zone is one model length deep. Both zones' widths are assumed to be equal to one hull length. To create a structured multi-block grid, the domain volume is divided into several sub volumes [10]. The grids at the free surface and near the model are refined in order to get better resolution of free-surface elevation in the area of importance and improve the boundary layer approximation [11]. According to Figure 2, the minimum grid spacing of the hull wall is $1 \times 10^{-3} L_{pp}$.

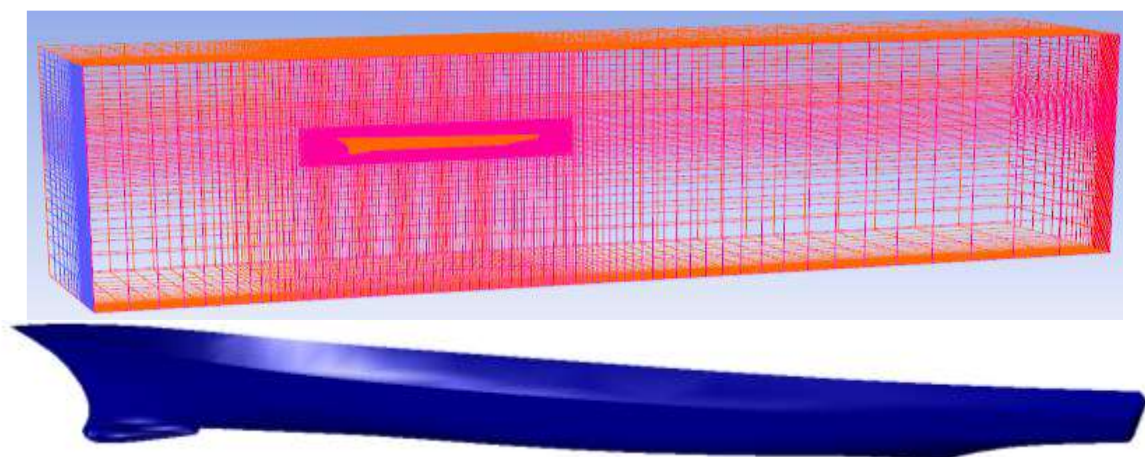


Figure 2 (A) Hull Domain Volume; (B) DTMB #5415 Model Geometry

While the flow velocity is taken to be equal to the experimental velocity of the model, the inlet and outlet boundary conditions upstream and downstream are taken as velocity inlet and pressure-outlet with an open channel, respectively. On all surfaces, a no-slip wall boundary requirement is applied. The symmetric plane invokes the symmetry requirement. Froude

numbers ranging from 0.1 to 0.45 were separately computed for various cases. The time step for this iteration is $\Delta t = 0.0001$ s [12].

A new model ship moving forward was simulated in the current investigation. According to ship sisters used by ITTC, a contemporary naval combatant surface ship was chosen as a recommended model for CFD validation for resistance and propulsion Table 2 and Figures 3 & 4 exhibit the main attributes and primary dimensions of this model. Water density is 999.8 kg/m^3 .

Table 3 Characteristics of Hull Model

Description	Units	Dimension s	Description	Units	Dimensions
Length between PP.	L_{pp} (m)	5.72	wetted Surface Area	S (m^2)	4.998
maximum breadth	B (m)	0.911	volume of displacement	∇ (m^3)	0.571
maximum draft	T (m)	0.235	block coefficient	C_B	0.465
Longitudinal CG	LCG (m)	2.258	kinematic Viscosity	(m^2/sec)	1.2845×10^{-6}

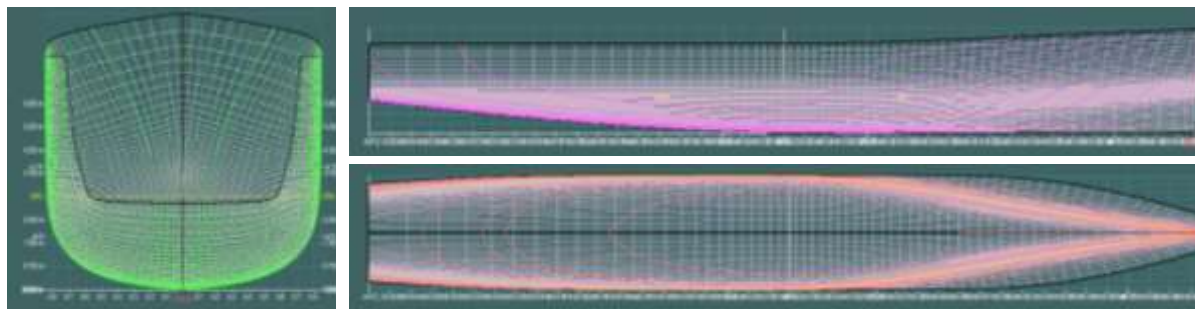


Figure 3 Lines Plan of The Hull Model

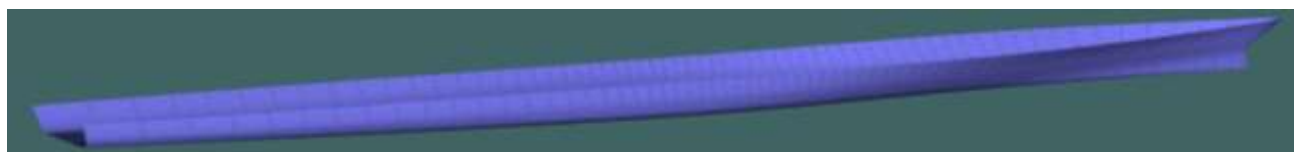


Figure 4 3 D Hull Model

The new model's resistance was calculated using different grid sizes using the same mesh, boundary conditions, and solution methodologies as the well-known DTMB #5415-51, and StarCCM+ results were compared to ensure competency with CFD theory [13].

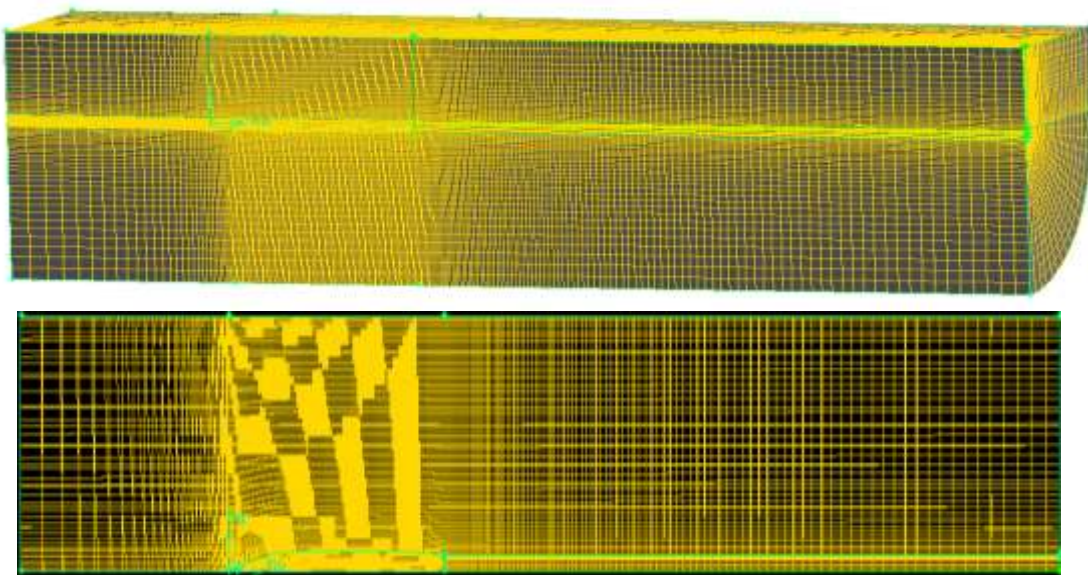


Figure 5 Hull Meshed Domain Volume

Three main processes are necessary for a CFD solution: processing, problem analysis, and post-processing of the outcomes [14]. In our work, the suitable mesh production is combined with the development of the hull geometry as the solution. Figures 5 & 6 depict the locations of the ship models in relation to the different limits of the solution domain as well as an overall view of the mesh surrounding the ship model [15]. For the flow computations, half of the new model according to the hull symmetry was employed. Three mesh sizes were investigated in this study; the total number of elements for coarse, medium, and fine grids, respectively, were 0.98, 1.35, and 1.82 M.

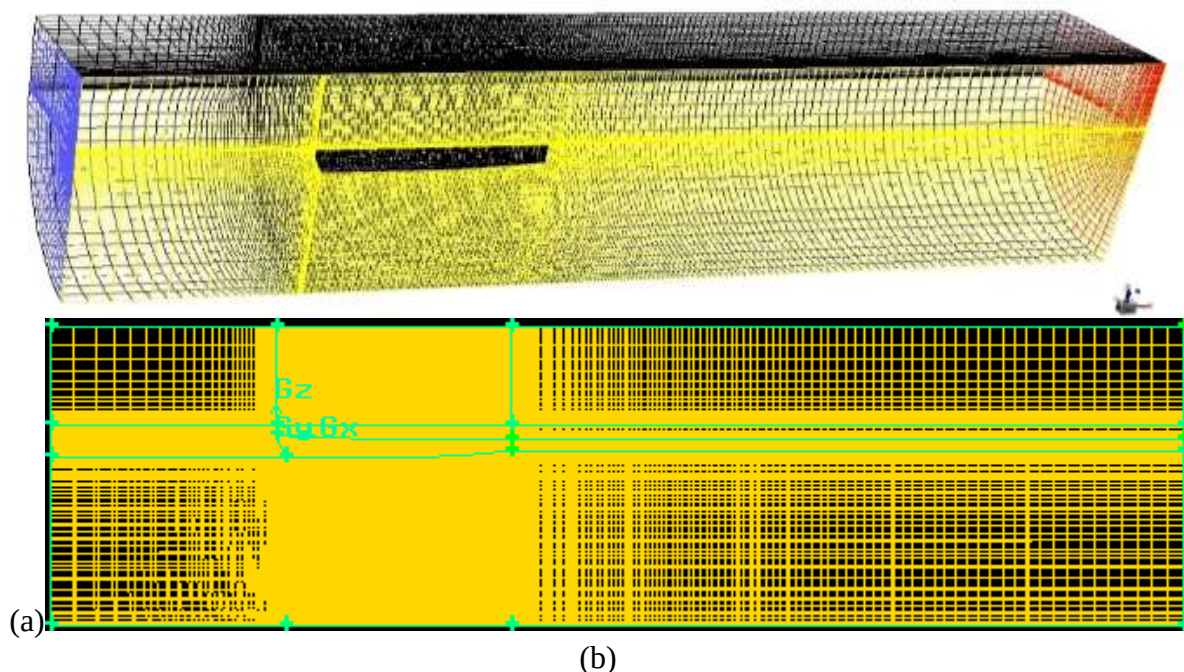


Figure 6 (A) Solution Domain (B) Overall View Of The Mesh Around The Model Ship

In this study, the Reynolds-averaged Navier-Stokes (RANS) equations are solved using the finite-volume method on hybrid structured grids using the StarCCM+. Gravity effects must be included in boundary conditions because the motion of the free-surface is controlled by

gravitational and inertial forces. The computations make use of the SST k-turbulence models with conventional coefficients. Monitoring the residuals of continuity, velocity, turbulence, volume fraction, and drag force allows for the assessment of the solution's convergence. The residual convergence criterion was taken as $1e^{-07}$ [16] [17].

Resistance Calculations

The DTMB #5415-51's total resistance was estimated. The experimental results from the ITTC and towing tank experiments, which are mentioned in Figure 7 [18] [19], were compared with the results for the three grids that are performed for different Froude numbers ranging from 0.10 to 0.45. The total and residual resistance at Fn 0.45 and the total and residual resistance coefficients (C_{TM} and C_R) for the estimated and experimental data are shown in Figures 8 & 9 respectively. The numerical calculation results of ship resistance with automatic running attitude adjustment are displayed in Table 3. The discrepancy between calculations and experiment findings is determined to be in good agreement, with an error of less than 4% where the difference is:

$$\Delta R_T\% = (R_T^{\text{StarCCM+}} - R_T^{\text{exp}}) / R_T^{\text{exp}}. \quad (5)$$

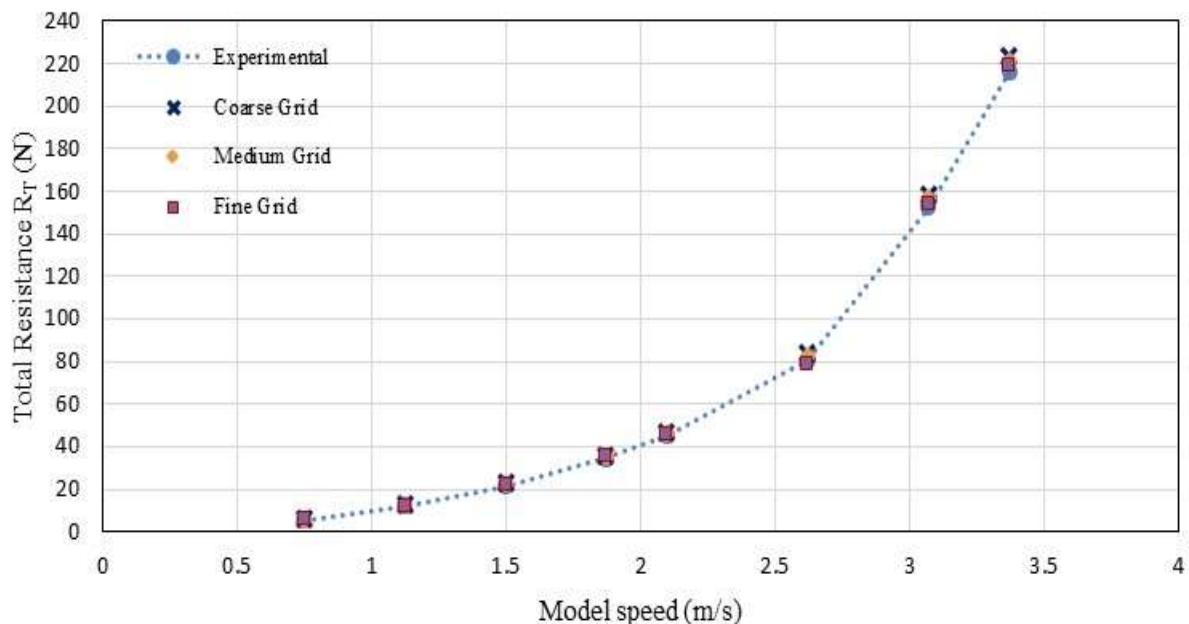


Figure 7 Resistance Curve for DTMB #5415 Comparing Three Different Sized Grids

Table 4 The Comparison of Experimental And CFD Results For DTMB #5415

F_n	R_T (N) (Experimental)	R_T (N) (C) (Simulation)	ΔR_T (%)	R_T (N)(M) (Simulation)	ΔR_T (%)	R_T (N) (F) (Simulation)	ΔR_T (%)
0.10	5.35	6.04	12.9	5.85	9.346	5.75	7.477
0.15	12.11	13.03	7.597	12.75	5.285	12	-1.32
0.20	21.58	22.65	4.958	22.18	2.78	22.1	2.41
0.25	34.68	36.02	3.864	35.52	2.422	35.3	1.845
0.28	45.18	46.59	3.121	46.32	2.523	46	1.837
0.35	80.66	83.65	3.707	82.77	2.616	79	-2.07
0.41	152.70	158.06	3.51	156.64	2.58	154	0.93
0.45	216.13	223.65	3.479	221.32	2.401	219	1.397

In light of the mesh generation method's applicability and the close agreement between the numerical predictions, it can be concluded that resistance prediction can be accomplished using the entire numerical scheme. The results for the coarse and medium grids are quite similar, and the difference between the medium and fine grids is comparatively bigger but still tolerably acceptable. As a result, the fine grid fits the computations best and produces the most relevant results. The same mesh generation technique may also be used to evaluate the hydrodynamic performances of the new combatant ship model using StarCCM+.

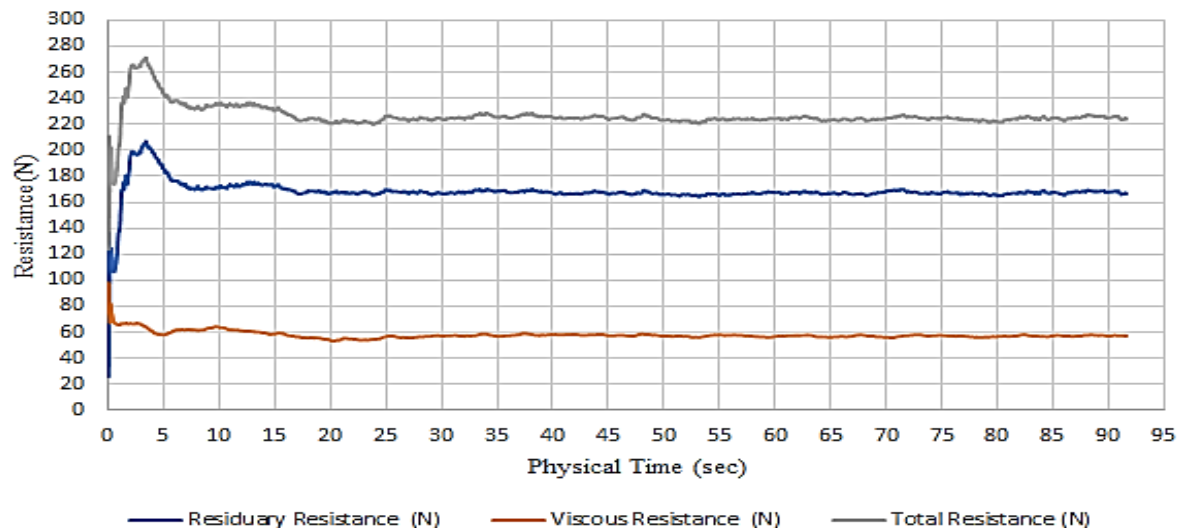


Figure 8 Resistance Plot for DTMB #5415 At Fn 0.45 At Fine Grid

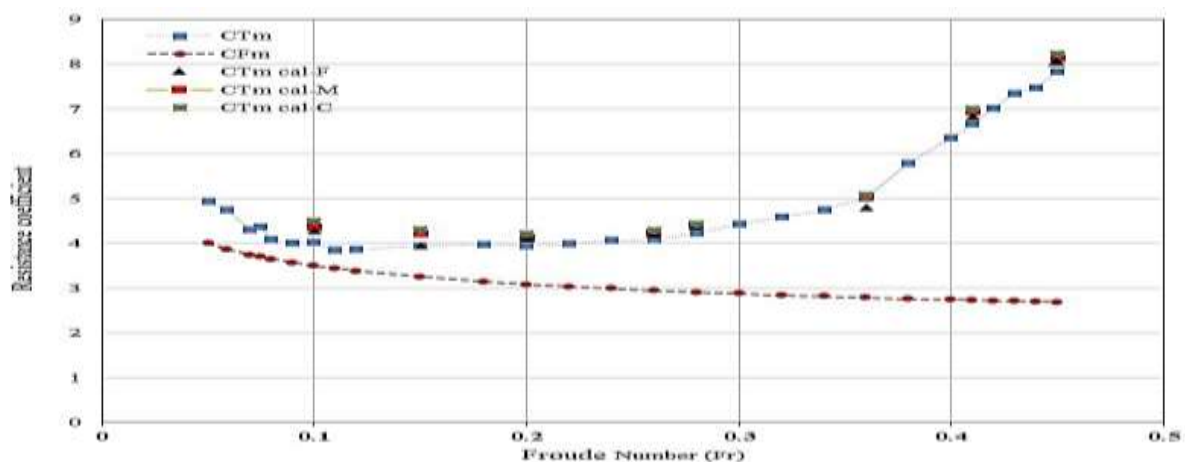


Figure 9 The Curves of Total and Residual Resistance Coefficient C_{TM} and C_R

A second early phase involved utilizing CFD to determine the new combatant ship model's resistance in comparison to Froude numbers ranging from 0.10 to 0.45 under identical conditions, as illustrated in Figure 10 where the values were compared also to Maxsurf results. For Three different mesh sizes were employed in the CFD mesh analysis to assess the hydrodynamic performances. To evaluate the convergence of the solution, the calculated drag on the hull was noted and plotted. The differences = (medium - coarse) / (fine - medium), where $0 < \epsilon < 1$ is minor and acceptable, are also in very close agreement with the results of the coarse, medium, and fine grids, as shown in table 4 [20].

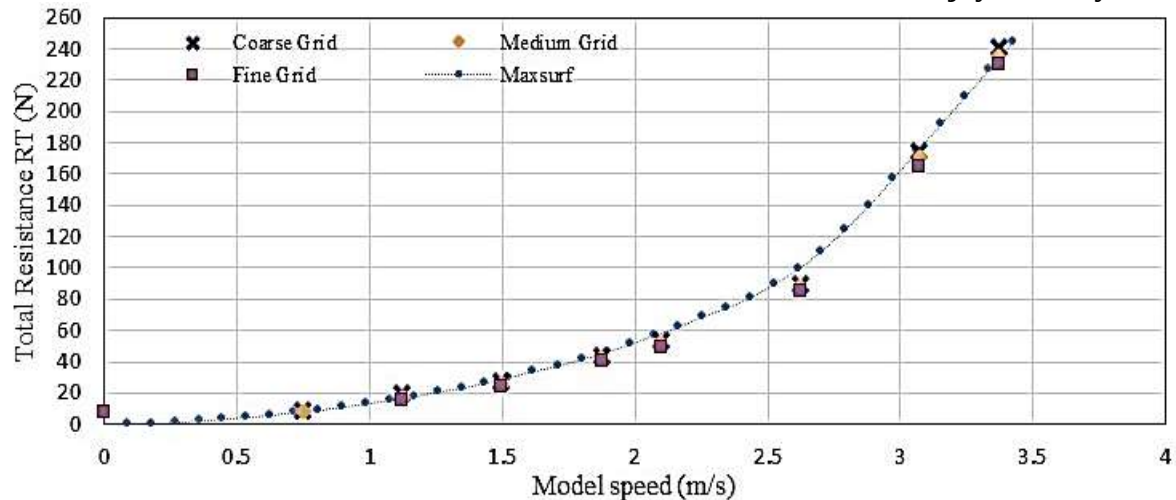


Figure 10 Resistance Curve for New Combatant Hull Model at Three Different Sized Grids

Table 5 The CFD Results for New Combatant Ship Model

F_n	0.10	0.15	0.20	0.25	0.28	0.35	0.41	0.45
R_T (N) (C)	9.12	19.65	27.1	44.12	53.89	89.34	175.15	240.65
R_T (N) (M)	8.56	17.47	26.32	42.15	51.66	87.45	170.35	235.26
R_T (N) (F)	7.32	15.25	24.68	40.12	49.33	85.55	165.19	229.75
Difference ϵ	0.45161	0.98198	0.47561	0.97044	0.95708	0.99474	0.93023	0.97822

In order to evaluate a mesh generation method under the same conditions and find a method to evaluate the combatant ship resistance calculations, the numerical predictions of DTMB #5415 and the new hull model are compared in this work as illustrated in Figure 11.

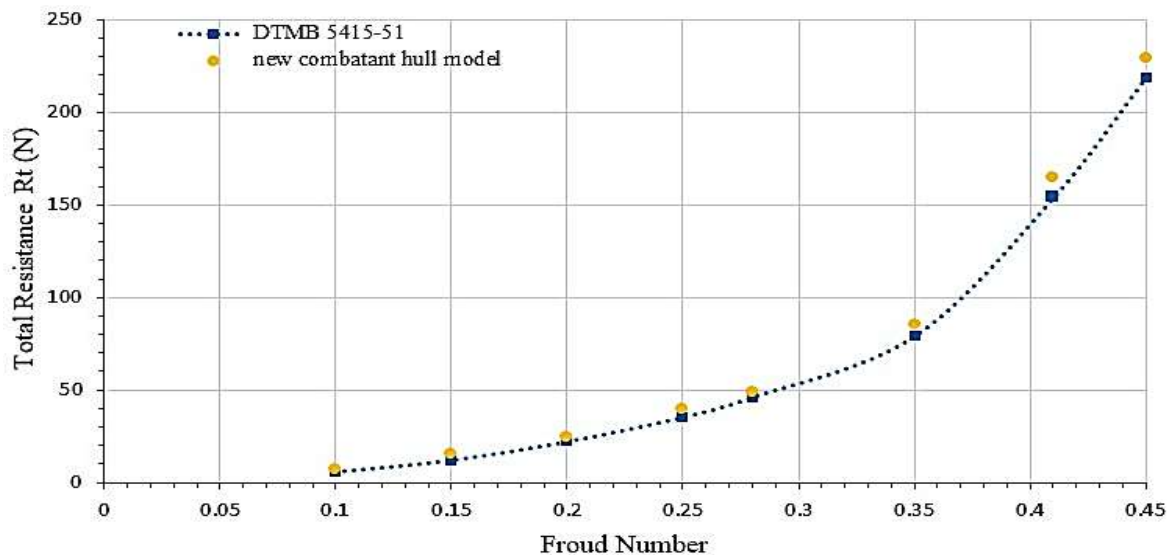


Figure 11 Resistance Results of DTMB #5415 And New Combatant Ship Model

The difference in total resistance between the two hulls is often acceptable for mesh investigation. The grid's results show excellent agreement. The expected resistance differs only slightly and by the same order of magnitude. As a result, the total numerical approach can be used to predict resistance.

Since convergent results are produced as the mesh size decreases, the fine grid is the one that will be used for the other hulls based on the grid results. It also fits the computation and displays the most relevant results. Finally, Figures 12, 13 and 14 illustrate the contours for the free surface wave and volume fraction (water) for the two hull models at various Froude numbers.

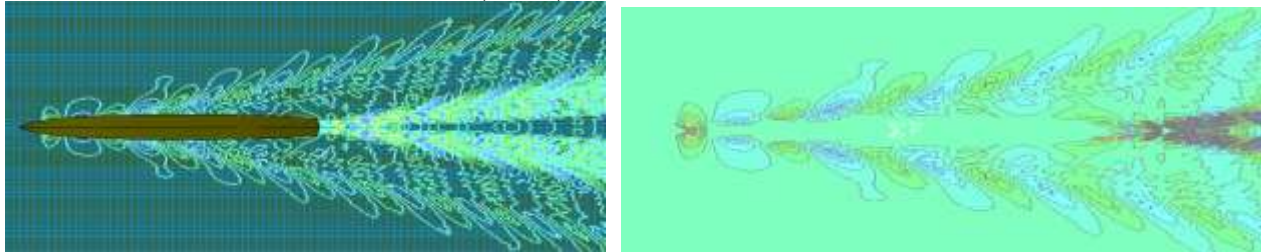


Figure 12 Free Surface Wave Contours for DTMB #5415 At $Fn = 0.40$ - Coarse Grid

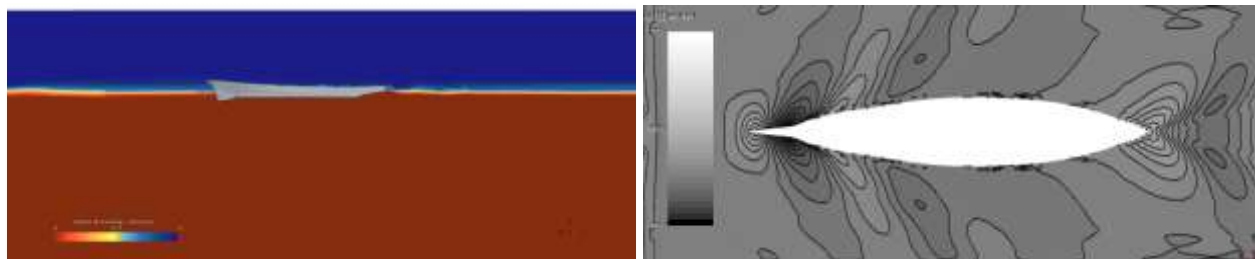


Figure 13 Volume Fraction for DTMB #5415 At $Fn = 0.45$ - Fine Grid

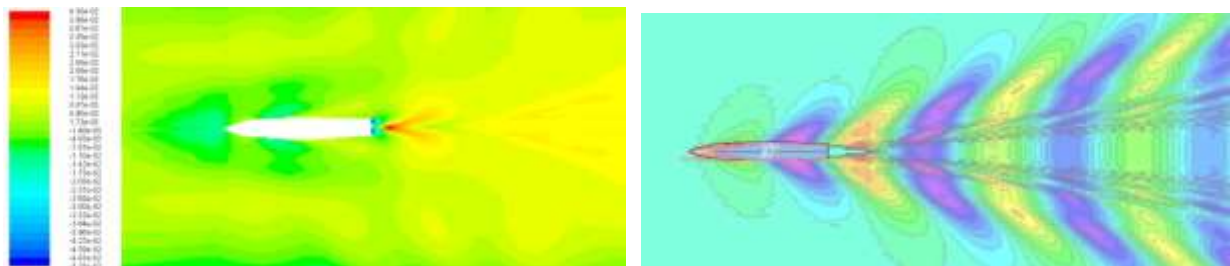


Figure 14 Free Surface Wave Contours for At $Fn=0.40$ - Medium Grid

Conclusion

An important portion of the current data set has been cross-validated using DTMB experimental results, with extremely positive results. The hull simulations for DTMB #5415 are within acceptable accuracy according to extensive comparisons between StarCCM+ calculations and the ITTC and experimental results. The method used is also acceptable, and the mesh generation method can be used to estimate the hydrodynamic performances for the model on the new combatant ship.

The fine grid is adapted to the calculation and shows the most appropriate results. The resistance calculations of the new combatant ship exhibited better performance than the DTMB #5415 under the same conditions to provide a promising test model for experimental investigation.

References

- [1] S. of the N.-W. W. Department of the navy, "Classification of Naval Ships and Craft," <https://doni.daps.dla.mil/>, 2006.
- [2] Z. rong ZHANG, H. LIU, S. ping ZHU, and F. ZHAO, "Application of CFD in ship

- engineering design practice and ship hydrodynamics,” *J. Hydrodyn.*, vol. 18, no. 3 SUPPL., pp. 315–322, Jul. 2006.
- [3] F. Stern, J. Longo, R. Penna, A. Olivieri, T. Ratcliffe, and H. Coleman, “INTERNATIONAL COLLABORATION ON BENCHMARK CFD VALIDATION DATA FOR SURFACE COMBATANT DTMB MODEL 5415*,” in *Twenty-Third Symposium on Naval Hydrodynamics*, 2001.
 - [4] S. Mancini, E. Begovic, A. H. Day, and A. Incecik, “Verification and validation of numerical modelling of DTMB 5415 roll decay,” *Ocean Eng.*, vol. 162, pp. 209–223, 2018.
 - [5] A. Elhadad, W. Y. Duan, and R. Deng, “Comparative Investigation of an Automated Oceanic Wave Surface Glider Robot Influence on Resistance Prediction Using CFD Method,” *Appl. Mech. Mater.*, vol. 710, pp. 91–97, Jan. 2015.
 - [6] F. R. Menter, “Two-equation eddy-viscosity turbulence models for engineering applications,” *AIAA J.*, vol. 32, no. 8, pp. 1598–1605, 1994.
 - [7] P. Akbarzadeh, P. Molana, and M. A. Badri, “Determining resistance coefficient for series 60 vessels using numerical and experimental modelling,” *Ships Offshore Struct.*, vol. 11, no. 8, pp. 874–879, 2016.
 - [8] M. Youssef, O. El Moctar, and T. E. Schellin, “Two Turbulence Models For URANS Simulations Of A Circular Cylinder Subject To Vortex-Induced Vibration,” *The 31st International Ocean and Polar Engineering Conference*. p. ISOPE-I-21-3221, 20-Jun-2021.
 - [9] A. Elhadad, W. Y. Duan, R. Deng, and H. Elhanfey, “Numerical analysis for resistance calculations of NPL as a floating hull for wave glider,” *Appl. Mech. Mater.*, vol. 619, pp. 38–43, 2014.
 - [10] L. Lazauskas, “Resistance and Squat of Surface Combatant DTMB Model 5415: Experiments and Predictions,” no. April, pp. 1–5, 2009.
 - [11] W. Rosemurgy, D. Edmund, K. Maki, and R. Beck, “A Method for Resistance Prediction in the Design Environment,” in *11th International Conference on Fast Sea Transportation, FAST 2011 - Proceedings*, 2011.
 - [12] A. Elhadad, “Resistance Evaluation for the Submerged Glider System using CFD Modelling,” *J. Adv. Res. Appl. Sci. Eng. Technol.*, vol. 3, no. 3, pp. 147–159, 2023.
 - [13] F. Pacuraru, “Validation of potential flow method for ship resistance prediction,” in *IOP Conference Series: Materials Science and Engineering*, 2019, vol. 591, no. 1.
 - [14] M. Diez *et al.*, “Multi-objective Hydrodynamic Optimization of the DTMB 5415 for Resistance and Seakeeping,” 2020, pp. 0–17.
 - [15] A. M. E. Ahmed and W. Duan, “Resistance and Seakeeping Investigation for Optimization of the Floating Hull of Wave Glider,” *World J. Eng. Technol.*, vol. 04, no. 03, pp. 235–249, 2016.
 - [16] A. Shi, M. Wu, B. Yang, X. Wang, and Z. Wang, “Resistance calculation and motions simulation for free surface ship based on CFD,” *Procedia Eng.*, vol. 31, pp. 68–74, 2012.
 - [17] A. Elhadad, W. Duan, and R. Deng, “A Computational Fluid Dynamics Method for Resistance Prediction of the Floating Hull of Wave Glider,” *Adv. Mater. Res.*, vol. 936, pp. 2114–2119, Dec. 2014.
 - [18] F. Zhao, S. P. Zhu, and Z. R. Zhang, “Numerical experiments of a benchmark hull based on a turbulent free-surface flow model,” *C. - Comput. Model. Eng. Sci.*, vol. 9, no. 3, pp. 273–285, 2005.
 - [19] K. Zelazny, “A method of calculation of ship resistance on calm water useful at preliminary stages of ship design,” *Zesz. Nauk.*, vol. 38, no. 110, pp. 125–130, 2014.
 - [20] S. Leksono; Muryadin; DKartikasari, “Comparative Study of Ship Resistance between



**4th International Research Conference on Applied Sciences and Engineering
Technology 2023 (4th IRCASET 2023)**

e-ISBN: 978-967-2426-49-3

Dorsett Hotel, Putrajaya, Malaysia

Model Test and Empirical Calculation of 60 GT Fishing Vessel,” *Int. J. Sci. Res.*, vol. 7, no. 12, pp. 1077–1082, 2018.

EXPERIMENTAL AND CFD RESISTANCE VALIDATION OF NAVAL COMBATANT DTMB 5415-51 MODEL

Alaaeldeen M. Elhadad^{1*}

Ahmed M. Abo El-Ela¹

¹Shipbuilding Engineering Department, Military Technical College, Cairo, Egypt

Abstract: The hull model of the surface combatant DTMB 5415 has been selected as a recommended benchmark naval ship for computational fluid dynamics (CFD) validation for resistance analysis. CFD is used to present a method for predicting surface combatant ship resistance. The resistance calculations of a 3 m DTMB 5415 hull model was compared using experimental and CFD techniques. The bare hull shape experiments were set up and carried out at the towing tank facilities at the Military Technical College's Hydrodynamic Laboratory in Egypt. A benchmark test will ensure that the equipment, procedures, and estimates of uncertainty are adequate. For the purpose of model validation, CFD calculations for a DTMB 5415-51 model are performed using three different mesh sizes for Froude numbers ranging from 0.10 to 0.40. For the same Froude numbers and free model conditions, results from towing tank experiments on the model's resistance, sinkage, and trim are presented. The numerical simulations are quantitatively consistent. The numerical results are compared in terms of wave field and resistance coefficients to determine the accuracy of the solution parameters. As evidenced by the resistance curves, the experimental investigation appears to provide very good agreement, indicating that the CFD model is capable of simulating the steady flow around a ship hull with acceptable accuracy and can thus be used as a supplement to laboratory model tests for ship design and ship hydrodynamic research.

Keywords: DTMB; towing tank; resistance; CFD

1. Introduction

A newly constructed ship's resistance characteristics are often calculated numerically using various methods or tested via model testing in a towing tank. For the CFD validation of resistance and propulsion, a large benchmark database is available. The 22nd International Towing Tank Conference's Resistance Committee released a report on the thorough tests that went into building this database [1]. About 1980, the US Navy came up with the DTMB 5415 hull form at the David Taylor Model Basin (DTMB) as a concept for a surface combatant ship with a sonar dome bow and a transom stem [2].

For the verification, validation/calibration, and accreditation of computational fluid dynamics (CFD) codes in order to realize simulation-based design, particularly for contemporary hull forms, there is an ongoing demand for more model-scale data. When complicated fluid flow is present, CFD techniques are very helpful for studying flow issues in resistance prediction. While towing tank tests provide better absolute accuracy. Also, they have the benefit of enabling hull shape alterations, which provides a comparative examination of findings to be conducted in a reasonably short amount of time and at a quite low cost [3].

All the benchmark experimental data used to validate the CFD results were also gathered using the same hull model. It is clear that different choices of the parameters produce different flow prediction. Due to these factors, not only for the global values but also for the local features of the flow, an experimental validation of the computed results is required [4].

The present study provides the test results for the resistance, sinkage and trim for combatant surface ship model DTMB 5415-51 to a scale of $\lambda = 51.2$ in at the Military Technical College towing tank (MTC-TT). The resistance, sinkage and trim data has been performed following the ITTC standards for model speed (V_m) from 0.01-2.0 m/s equivalent to Froude numbers (Fr) 0.05-0.40. The results are repeated ten times for $V_m = 0.5, 1.0, 1.5$ and 2.00 m/s for checking accuracy. The results are discussed with regard to the data trends and compared to the CFD results.

This paper's requirement is to determine an accurate numerical prediction of hull model resistance, sinkage, and trim, as well as the flow field surrounding the hull model on various Froude number scenarios [5]. CFD results are as a primary step for production of a validated ship model then compared with experimental investigation in towing tank for model verification. The paper aim to compare the experimental results in MTC-TT with already documented results on DTMB 5415 3m Bench mark to ensure adequacy of laboratory setup ad procedures [6]. Results of the extensive experimental campaign are presented forming a benchmark for numerical methods validation.

2. Numerical Modelling

2.1 Hull Geometry

Many experimental and CFD test results for the surface combatant DTMB model 5415 are public and have been addressed at numerous workshops and conferences, including the Gothenburg, Tokyo, and SIMMAN workshops. Figure 1 and Table 1 provide a 3-D representation of the hull and the key details, which, when compared to the model tests from the David Taylor Model Basin in Washington, D.C., are accurate. The bare hull condition is used for computations and experimentation [7].

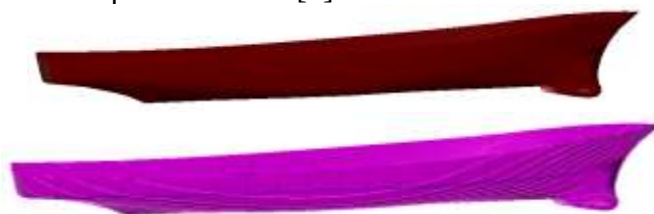


Figure 15 DTMB #5415 Model Geometry

Table 6 Main particulars of DTMB #5415 (ITTC, 2005)

L_{pp} (m)	2.78	Δ (kg)	63.5
B (m)	0.403	LCG (m)	1.375
D (m)	0.244	Wetted Surface Area (S) (m ²)	1.313
T (m)	0.120	Block Coefficient (C_B)	0.506

Figure 2 (a) shows the body plan of DTMB model which is believed to be one of the best choices to simulate turbulence flow around ship hull. It is usually regarded as a preferred test sample to validate a new numerical method on this combatant hull [8] [9]. The grid structure

and the profile of the combatant ship model corresponding to a scale of 51.10 used for CFD calculations are illustrated in figure 2 (b).

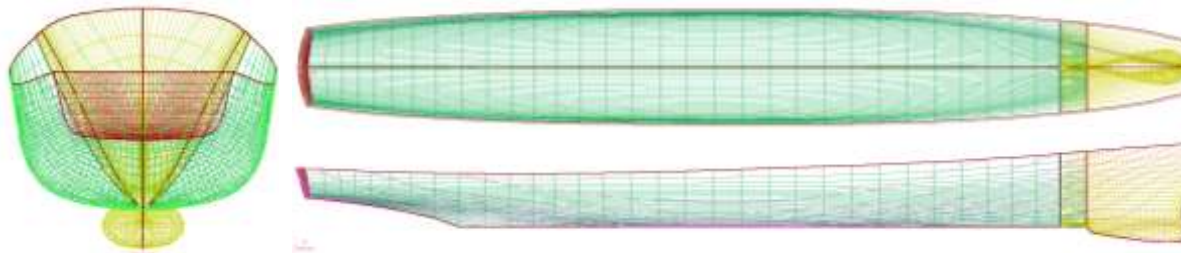


Figure 16(A) Body Plan (B) Numerical Grid of DTMB 5415 Hull Surface

2.2 Mathematical Model and Governing Equations

The continuity equation, in conjunction with the Reynolds-averaged Navier-Stokes (RANS) equations, can be used to represent the governing equations of an incompressible, two-phase (air and water), and viscous turbulent flow field. The SST $k-\omega$ Model was used to calculate the turbulent viscosity; further information on the model formulation and turbulence model equations is provided by Menter [10]. The VOF method is utilised to simulate multiphase free-surface flow.

The fluid domain is partitioned into a finite number of cells, and a discretionary process into algebraic form to be solved then converts the governing equations for fluid flow. A simple algorithm was used to achieve the coupling of the pressure and velocity fields. The computation is carried out in Fluent, a multipurpose CFD software that was recently used to simulate ship flow [11]. In this study, we demonstrate that the calculation for ship resistance with varying Froude number can be undertaken in the Fluent as well.

2.3 Computational Domain, Boundaries and Mesh Configuration

The rectangular hull-meshed domain is separated into the water zone and the air zone. All analytical and numerical solutions must be applicable to the given boundary and initial conditions. These requirements need to be specified in accordance with the flow characteristics. The resistance behavior of the DTMB 5415 hull in deep water was predicted computationally in the current work. To reduce computational time, only half of the model conforming to the hull symmetry was modelled. Figure 3 displays the main hull's boundary conditions. While the flow velocity is taken to be equal to the experimental velocity of the model, the inlet and outlet boundary conditions upstream and downstream are taken as velocity inlet and pressure-outlet with an open channel, respectively. No-slip walls with zero normal and tangential velocities were used to create ship limits. As a result, both kinematic boundary condition and no-slip condition were satisfied on the hull surfaces [12]. Symmetry condition is invoked on the symmetric plane.

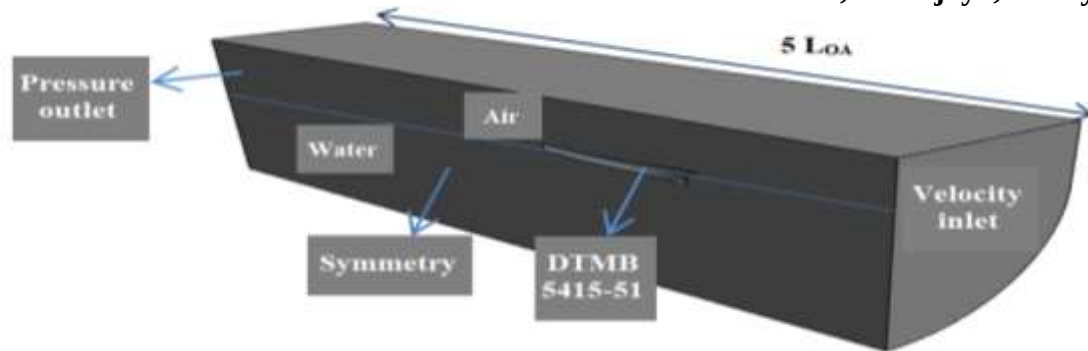


Figure 17 Computational Domain and Boundary Conditions Of The DTMB 5415

It should be observed that the main hull and the corresponding hulls with other scales are geometrically identical. The x, y, and z-axes are oriented with x-axis pointing towards the bow, y-axis to portside and positive z upwards and negative z downwards directions. Table 2 also provides the computational domain's dimensions.

Table 7 Computational Domain Dimensions

Upstream	1.0 L _{OA}	front of the bow
Down stream	3.0 L _{OA}	from the stern
Top	0.5 L _{OA}	depth of air zone
Bottom	1.0 L _{OA}	depth of water zone
Transverse	1.0 L _{OA}	width of both zones

To create a structured multi-block grid, the domain volume is divided into several sub-volumes [13]. In order to acquire greater resolution of the free-surface elevation in the important area and improve the boundary layer approximation, the grids at the free surface and close to the model are refined. According to figure 4, the minimum grid spacing away from the hull wall is $1 \times 10^{-3} L_{PP}$.

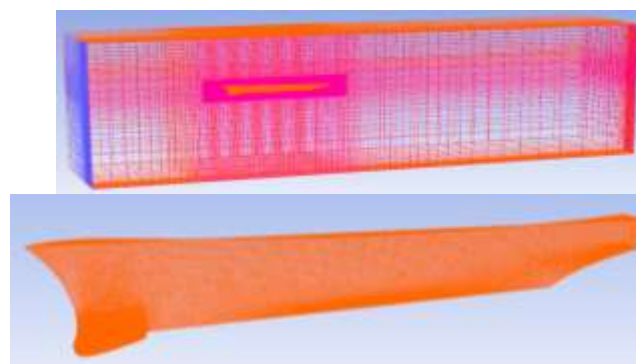


Figure 18 (A) Hull Domain Volume in Gambit; (B) DTMB #5415 Model Cells

Three main processes are necessary for a CFD solution: processing, problem analysis, and post-processing of the outcomes [14]. In our work, the suitable mesh production in the Gambit programme is combined with the development of the hull geometry as the solution. Figure 5 depicts the locations of the ship model relative to the various solution domain boundaries as well as a general perspective of the mesh surrounding the ship model. Over the entire solution domain, a hybrid structured/unstructured mesh with around 1.15 million cells was produced [15]. The computational domain was discretized using hexahedral components. Local grid refinements were used close to the free surface and around the hull. In this investigation, three

different mesh sizes were examined with a total element count of 1.15 M, 1.32 M and 1.45 M for coarse, medium and fine grid respectively.

The calculations for this work were performed using the Fluent programme, which employs a hybrid-structured grid to solve the RANS equations using a finite-volume method [16]. Gravity effects must be included in boundary conditions because gravitational and inertial forces control the motion of the free surface. The calculations make use of the SST $k-\omega$ turbulence models with standard coefficients. Monitoring the residuals of continuity, velocity, turbulence, volume fraction, and drag force allows for the assessment of the solution's convergence. The assumed value for the residual convergence criterion is $1E-07$ [17]. Froude numbers ranging from 0.1 to 0.45 were separately computed for various scenarios. The Iteration's time step is $\Delta t = 0.0001s$.

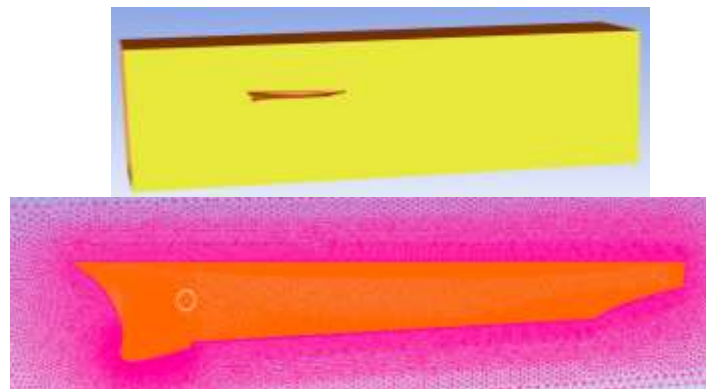


Figure 19 (A) Solution Domain (B) Overall View of The Mesh Around The Model Ship

3. Experimental Setup

3.1 Description of The Facility and Equipment

The experiments have been carried out in the Hydrodynamic Laboratory at the Military Technical College, Cairo. The towing tank (MTC-TT) has the geometrical specifications; 45 m long, 5 m wide and has a water depth 3.8 m and technical specifications, which include a J-frame carriage with speed to 2 m/sec and carriage max payload of 50 kg. It is equipped with a complete wave maker system that includes eight multi-flap paddles designed and built by Edinburgh Designs Ltd. as seen in figure 6. To stabilise operation and achieve the appropriate transfer function, the wavemaker uses force feedback mode in the electronic control system. Moreover, the wave maker incorporates absorption facilities to eliminate the impacts of reflected waves [18].

The model was placed in the tank's centre, both longitudinally and transversely. To measure the elevation of the water's surface, a resistance wave gauge controller with eight channels and eight probes was set up. With one resistance probe placed 15 cm from the wave maker and one ultrasonic wave probe close to the model, it was observed and recorded. The mass density and viscosity of water are calculated using the ITTC procedure [19]. Before conducting model testing, the dynamometer was calibrated using masses in accordance with the ITTC Protocol [20].

Systems for signal conditioning and data collection are installed on the driving carriage. To minimize errors in the extension of the ship, the measurement of the model resistance must be exceedingly precise. The most crucial prerequisite for measuring the resistance of a ship model is consistency of speed for the towing carriage, together with the precision of the resistance dynamometer made by Edinburgh Designs Ltd. in conjunction with a model clamp.

On a Kempf and Remmers accuracy test platform, the load cell, signal conditioner, and carriage PC AD card are statically calibrated to ascertain the voltage-mass relationship. To lessen the sensitivity of the signal to noise, amplified analogue voltages are transformed into frequencies (3000 ± 2500 Hz) for transmission to the AD card [21]. The carriage is driven at the desired constant speed during each test run, and measurements of the speed and model resistance are made to help interpret the data afterwards.

With a 5-camera system properly positioned above the water's surface, the Qualisys video motion capture system was used to measure the model's six degrees of freedom of motion. It is comprised of seven infrared reflectors strategically placed on the vessel (illuminated balls). The coordinates of the balls in 3D space are registered by five cameras suitably positioned near the vessel and the six-degree of freedom motions are calculated and output in real time.

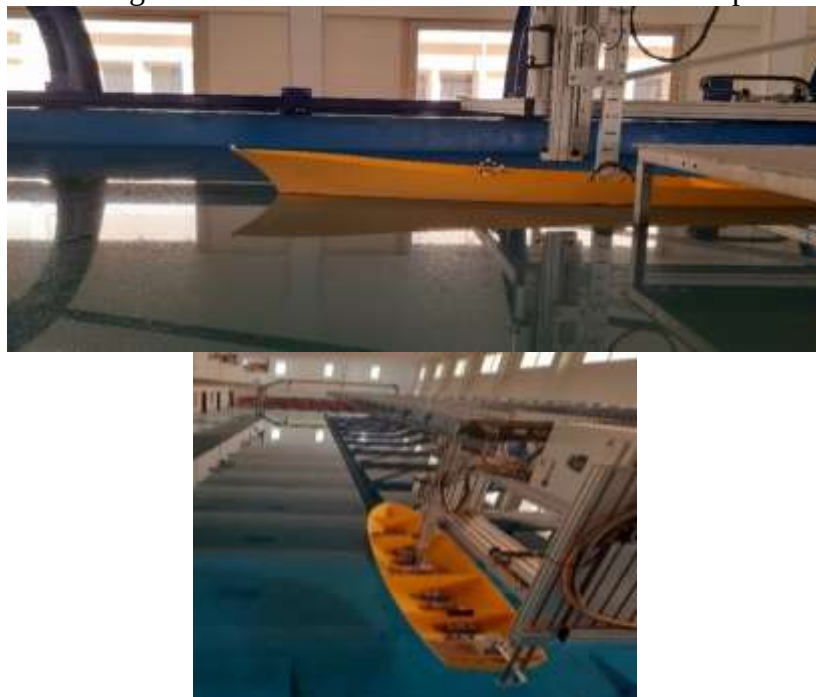


Figure 20 (A) DTMB #5415 Model and Experimental Set Up; (B) MTC- Towing Tank

3.2 Construction of the Model

The International Towing Tank (ITTC) has chosen the DTMB 5415 hull shape model as a suggested benchmark naval ship for CFD validation for a resistance study [22]. The model presents a transom stern and a bow bulbous of peculiar shape that allows the sonar lodging. The length between perpendiculars (L_{pp}) of the general tested DTMB model is 5.72 m, which translates to a scale of 24.8. All of the testing was done in bare hull conditions. In our tests, one identically shaped model was tested. It was constructed in the model factory of Edinburgh Designs Ltd. The MTC-TT model is constructed from fiberglass. The main components of the geosim model and the hull lines plan are shown in table 3 and figure 7, respectively. The 1:51 model's internal geometry matched that of the ITTC presentation exactly. The four watertight bulkheads have been built on the model, and figure 8 shows where they are located in relation to ship scale.

The model adopted a row of cylindrical studs with a 3 mm height and 3 mm diameter, spaced 30 mm apart, and fitted them 60 mm behind the bow profile in order to promote turbulent flow.

The size and spacing of the studs follows [23]. In the following discussion, as well as in the introduction, we refer to the model as DTMB 5415-51.

Tolerances on the model must be kept to a minimum. The waterline at the design draught was free of trim and sinkage, and the surface was correctly completed and ballasted to the required displacement. It was attached to the resistance dynamometer of the towing carriage according to ITTC procedure [24].

Table 8 Main particulars of national US combatant Ship DTMB 5415

Particulars	Ship	Model 24	Model 51
L_{OA} (m)	153.300	6.167	3.0
L_{PP} (m)	142.200	5.72	2.780
B_{OA} (m)	20.540	0.826	0.403
B_{WL} (m)	19.082	0.767	0.375
D (m)	12.470	0.502	0.244
T (m)	6.150	0.248	0.120
V (m ³)	8424.4	0.549	0.0635
Δ (t,kg)	8636	0.549	63.5
KM (m)	9.493	0.382	0.186
KG (m)	7.555	0.304	0.148
GM (m)	1.938	0.078	0.038
LCG (m)	70.137	2.884	1.375
Hull Coefficient			
C_B	0.505	L_{PP}/B	7.530
C_P	0.616	B/T	3.091
C_M	0.815	C_W	0.778

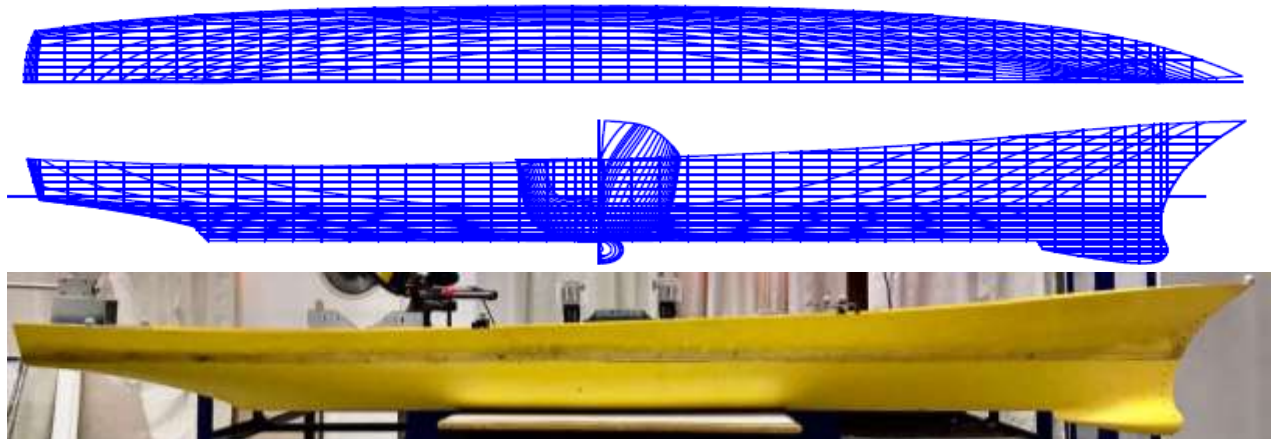


Figure 21 DTMB #5415 Geometry, Photo and Lines Plan

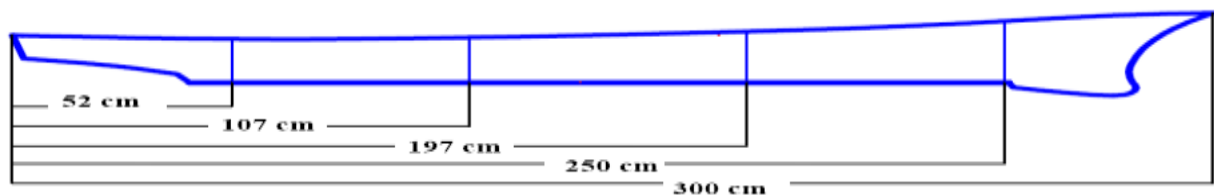




Figure 22 Model Subdivision—Dimensions Are in Ship Scale

3.3 Description of Test Conditions and Test Trials - Tests Design

The MTC-TT test used the standard towing tank tests for resistance, sinkage, and trim, and it was conducted as usual [25]. Accuracy assessments are essential for experiments that use processes. An uncertainty analysis was created before to the testing in order to assess the level of accuracy that might be expected from the outcomes. The AIAA rules state that uncertainty analysis is a well-established method for determining if experimental results are accurate and that it should be required at all stages of experimentation [26]. It is helpful for reviewing data obtained for comparing data from many studies throughout the planning and development phases of research. As part of the internal development and application of uncertainty analysis processes and the assessment of scale effects and facility bias, conventional uncertainty assessment procedures are used to determine the data quality [27].

Resistance measurements have been made in free-model settings with Froude numbers ranging from $Fr = 0.05$ to 0.40 , as well as sinkage and trim. The forces have been measured using load cells, while sinkage and trim have been determined using a mechanism that converts encoder-recorded translations into rotations. A forward-perpendicular and design waterplane intersection has been chosen as the origin of a Cartesian coordinate system. The (x, y, and z) axes are oriented in that order (downstream, transverse, and upward) respectively [28].

Towing tank water temperature is measured regularly to allow water viscosity to be estimated for the calculation of Reynolds number and hence friction coefficient at the model mid draft using a digital thermometer as shown in figure 9. The water data will allow the calculation of viscosity and density, and thus the Reynolds number, for each run. Tests were carried out at water kinematic viscosity $\nu_M = 1.16 \times 10^{-6} \text{ m}^2/\text{s}$ and density 999.2 kg/m^3 . The towing force in (kg) is converted to Newtons (N) by multiplication with $g = 9.81 \text{ m/s}^2$. The resistance values of the ship model are estimated for calm water conditions without the appendages (the rudder and the propeller) [29].

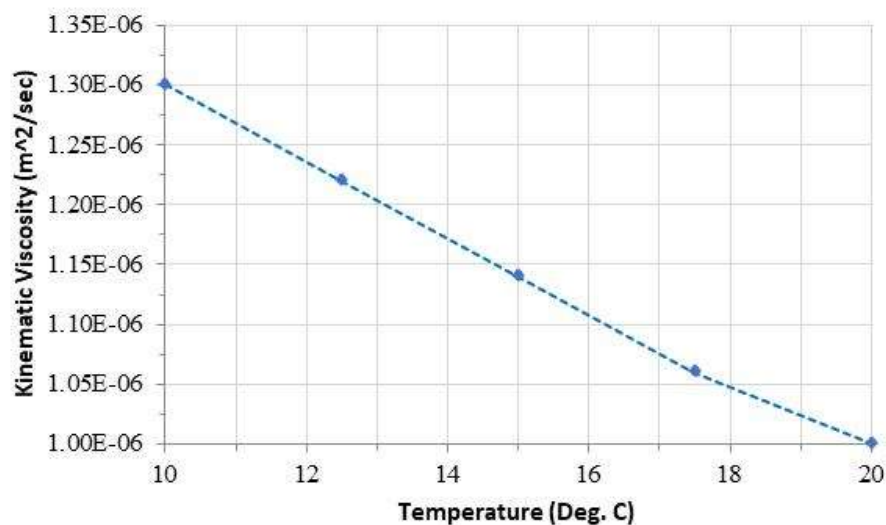


Figure 23 Viscosity Vs Temperature in Fresh Water

3.4 Setup and Experimental Procedure

The ITTC standards for model speed V_m between 0.1 and 2 m/sec were followed in performing the resistance test. Three experiments were conducted under identical circumstances on different days. To establish the precision limit, ten repeated readings have been taken at $V_m = 0.5, 1.0, 1.5$, and 2 m/sec. A load cell with a maximum load of 100 N has been used to assess the total model resistance. Before converting from voltage to frequency (V/F), the cell signal was boosted in order to lessen its vulnerability to transmission line noise. As a result, it was sent at a frequency of 3000 ± 2500 Hz, translating to a voltage range of ± 5 V, which was then transformed (F/V) into an output voltage for the purpose of acquisition by a 12-bit acquisition board mounted on a computer.

Using the same tools that would eventually be mounted on the carriage for the trials, the load cell calibration was completed [30]. Tachometric equipment has been used to measure the velocity, which the carriage control system has kept at the predetermined value. For every metre of carriage displacement, a tachometric wheel generates 1000 pulses on an optical encoder, yielding a spatial resolution of 0.001 m. The velocity in mm s^{-1} was calculated from the pulses that a 16-bit binary counter counted over a 1 s time. This number is transferred to the acquisition board and shown on the carriage control panel [27].

At the same Froude numbers and resistance, sinkage and trim have been determined by measuring the displacements of two points that are respectively close to the fore and rear perpendiculars [31]. Two angular potentiometers coupled to two pantographs have been used to measure the displacements of the two spots. Due to the very small trim angle, the relationship between fore-and-aft displacements and the trim angle has been linearized [32].

Rotative potentiometers were used to calculate sinkage and trim by measuring fore and aft displacements. Displacement measurements were obtained by converting vertical displacements to angular displacements using weighted, mechanical parallelograms. Potentiometers, signal conditioners, and the carriage PC AD card are statically calibrated to establish the voltage-displacement connection. To obtain the data, 300 discrete samples were collected at a rate of 30 Hz over the course of 10 seconds. A low-pass filter operating at 10 Hz filters the data. The first-order relation has calculated the trim angle by:

$$\theta \cong \frac{\Delta i_{FM} - \Delta i_{AM}}{L_{MM}} \quad (1)$$

where Δi_{FM} and Δi_{AM} are the displacement of the point that are near the fore and aft perpendicular respectively and L_{MM} is the distance between these two points [33]. The displacements of the two points, corresponding to the fore and aft perpendicular, have been obtained knowing their distances from the measurement points by relations:

$$\frac{\Delta i_{FM} - \Delta i_{AM}}{L_{MM}} = \frac{\Delta i_{FP} - \Delta i_{AM}}{L_{MM} + D} \Rightarrow \Delta i_{FP} = \Delta i_{AM} + \frac{L_{MM} + D}{L_{MM}} (\Delta i_{FM} - \Delta i_{AM}) \quad (2)$$

$$\frac{\Delta i_{FM} - \Delta i_{AM}}{L_{MM}} = \frac{\Delta i_{AM} - \Delta i_{AP}}{L_{PP} - L_{MM} - D} \Rightarrow \Delta i_{AP} = \Delta i_{AM} - \frac{L_{PP} - L_{MM} - D}{L_{MM}} (\Delta i_{FM} - \Delta i_{AM}) \quad (3)$$

where D is the separation between the measuring point's front and the front perpendicular. The following relationship has been used to determine sinkage:

$$\Delta i = \frac{1}{2} (\Delta i_{FP} + \Delta i_{AP}) = \left(1 - \frac{L_{PP} - 2D}{2L_{MM}}\right) \Delta i_{FM} + \left(\frac{L_{PP} - 2D}{2L_{MM}}\right) \Delta i_{AM} \quad (4)$$

4. Results and Discussion

4.1 Numerical Calculations

In the first preliminary step, the DTMB 5415-51 model's calm water resistance was tested using CFD against model speeds ranging from 0.25 to 2.3 m/sec, or Fn 0.05 to 0.45. Three different

mesh sizes were employed to estimate the hydrodynamic performances for the CFD mesh investigation. To assess the convergence of the solution as depicted in figure 10, the calculated drag on the hull was measured, plotted, and compared with the Maxsurf results. The findings for the coarse, medium and fine grid results are extremely similar and the difference $\varepsilon = (\text{medium} - \text{coarse}) / (\text{fine} - \text{medium})$, where $0 < \varepsilon < 1$ is small and acceptable as shown in table 4 [16] [34].

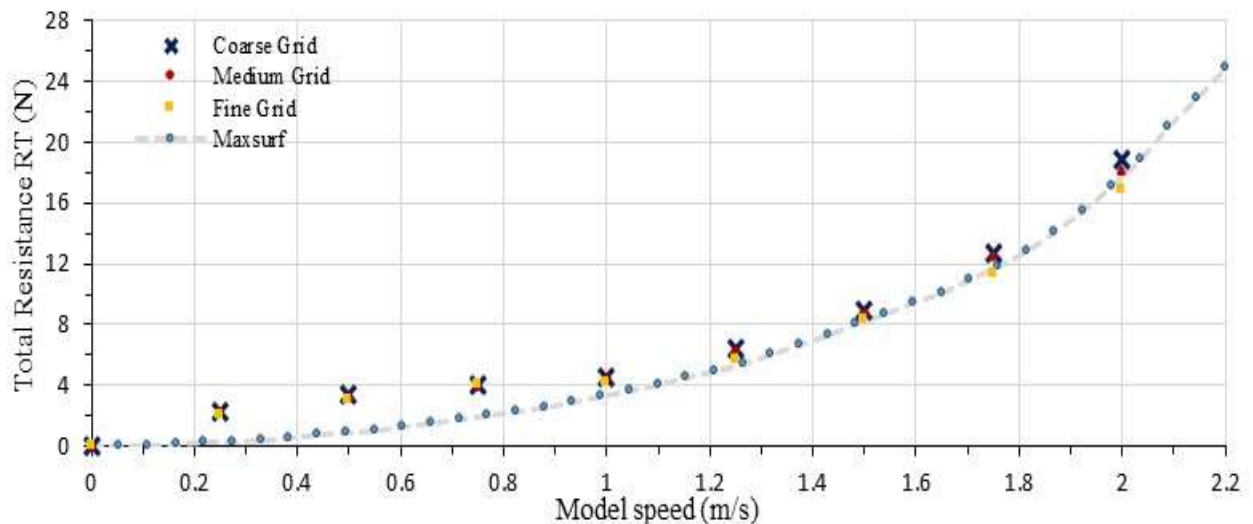


Figure 24 Resistance Curve for DTMB #5415-51 Comparing Three Different Sized Grids with Maxsurf

Table 9 The CFD Results for DTMB 5415-51 Model

V _m (m/sec)	0.25	0.5	0.75	1	1.25	1.5	1.75	2
R _T (N)	2.2635	3.3910	4.0081	4.5142	6.3920	8.9215	12.774	18.843
(C)	21	23	25	48	25	42	02	54
R _T (N)(M)	2.1524	3.2541	4.0005	4.4521	6.2541	8.7441	12.541	18.124
	87	65	2	3	8	21	52	84
R _T (N) (F)	1.9954	2.9874	3.9854	4.1054	5.6524	8.3125	11.332	16.845
	22	45	12	81	81	84	54	21
Differen ce ε	0.7069	0.5131	0.5033	0.1791	0.2290	0.4111	0.1923	0.5616
	29	15	76	96	93	37	12	48

Figure 11 illustrates the total resistance coefficient for the three grids with speed. According to the bare hull results, the mesh generation approach may be used to estimate hydrodynamic performances using CFD, and convergence of the numerical predictions demonstrates that the overall strategy is appropriate for resistance prediction. The results for the coarse and medium grids are quite similar, and the difference between the medium and fine grids is comparatively bigger but still tolerably tiny. Therefore, the fine grid provides the best match for the calculations and displays the outcomes, which are most appropriate.

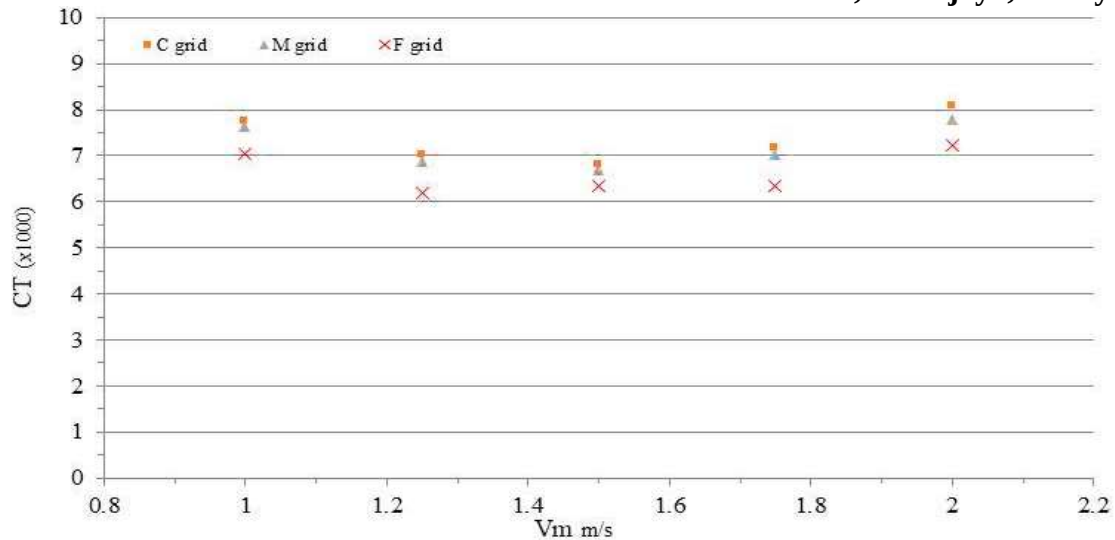


Figure 25 The Curves of Total Resistance Coefficient CT_M For Different Grids

Figure 12 shows the time histories of the total resistance, at Fr 0.30 for the 125 seconds that the simulation was run. It is possible to see how the graphs behave oscillatorily as a result of the free surface. The resistance values shown in this figure must be multiplied by two since, as previously mentioned, only half of the ship was simulated and the symmetry requirement was applied. The frictional and residuary resistances, into which the total resistance is divided, can also be provided by the CFD code [35]. While the frictional resistance is largely steady, the pressure resistance curve can be seen to oscillate. The oscillations of the pressure resistance value that can be attributed to the effect of the free surface are transmitted to the values of the total resistance.

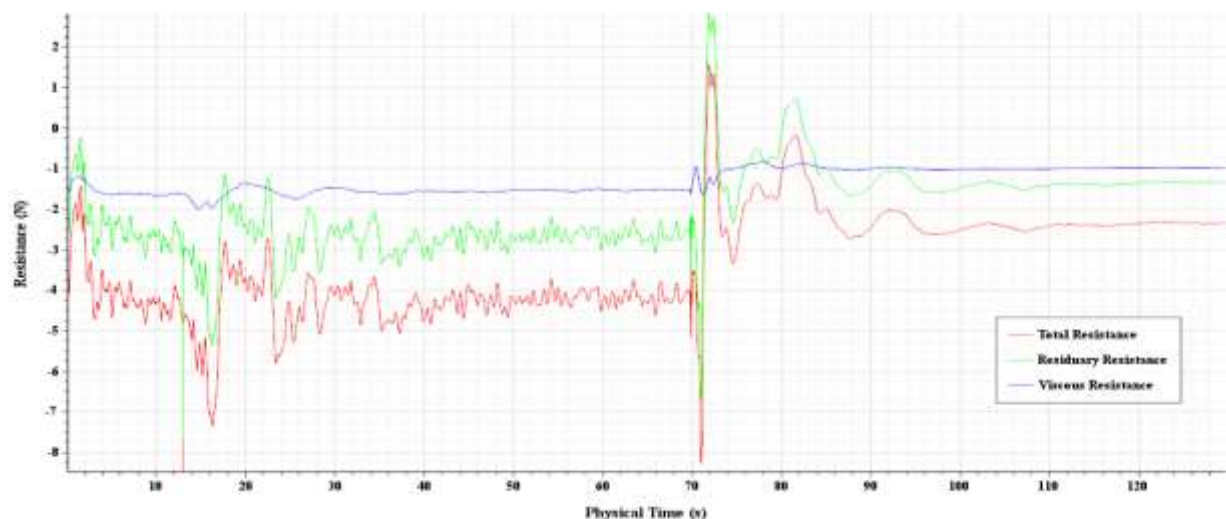


Figure 26 Time History of Resistance as Generated in CFD At $Fr=0.38$

Finally, figures for the hull model's free surface wave and volume fraction (water) contours at various Froude numbers are presented in figures 13 & 14 respectively.

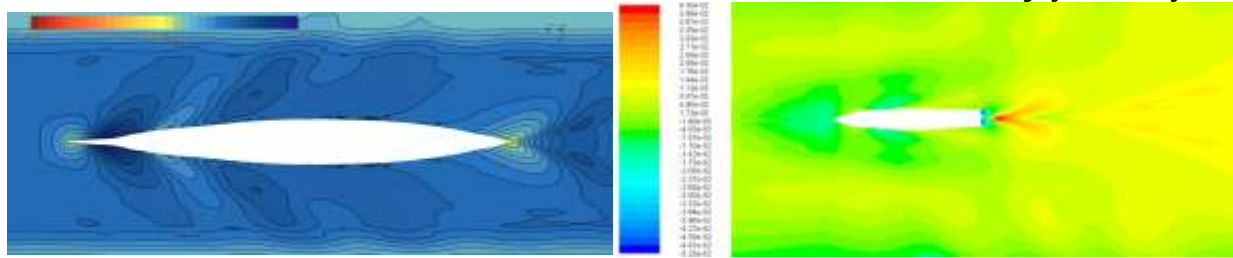


Figure 27 Wave Pattern for Bare Hull At $Fn = 0.40$ - Coarse Grid



Figure 28 Free Surface Wave Contours for At $Fn=0.40$ - Medium Grid

4.2 Experimental Results; Resistance, Sinkage And Trim

This section analyses and explains the experimental data that was gathered. It is believed that the experimental data comparison is sufficient to provide a thorough understanding of the fundamental traits of the difficult methodologies under investigation. Resistance benchmark experiments were conducted between July and November 2022 to create reference test settings that would allow for the comparison of experimental results from various tests. The benchmark testing outcomes in calm water were examined and the numerical outcomes were contrasted. Plotting resistance, sinkage, and trim versus speed during the testing is crucial for ensuring that the findings are consistent and repeatable. Target speed, measured speed, resistance, sinkage, trim, and water temperature should all be recorded for each run. Variation of model resistance results (for 10 repeats) should normally lie in a range $\pm 0.5 - 2\%$ of the mean.

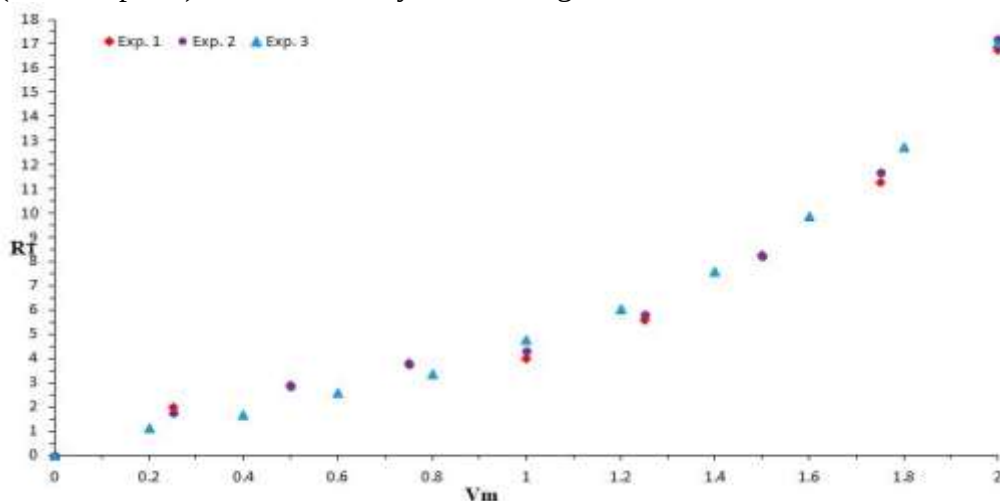


Figure 29 The Total Resistance Experiment Data for DTMB 5415-51

The total resistance curve as a function of carriage speed is shown in figure 15. As prescribed by the ITTC comparative tests, there would be used 10 repeat tests on 4 different days at measured speed $V_m = 0.5, 1.0, 1.5$ and 2.00 m/s, corresponding respectively to $Fr = 0.10, 0.24, 0.34$ and 0.39 , to perform statistical analysis as demonstrated in figure 16. Before doing any

statistical analysis, the total resistance data in the towing tank are adjusted to the nominal speed and converted to the nominal fresh water temperature of 15 degrees Celsius.

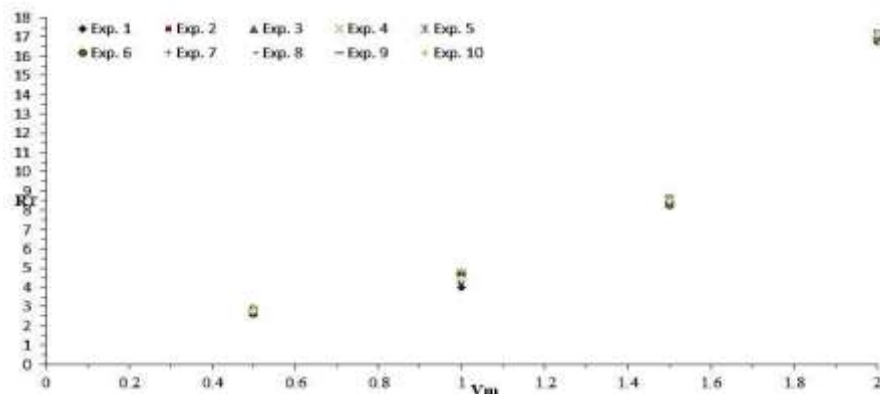


Figure 30 The Total Repeat Resistance Tests for DTMB 5415-51

The means of total resistance from those repeat tests are given in table 5. Such means can be regarded as the most accurate in each resistance test in the towing tank. The experimental standard deviation (StDev) of tests is also presented. These standard deviations can be used to calculate the measurement's repeatability and uncertainty.

Table 10 Statistical Analysis Resistance Value for Hull Model

Fr	Mean	StDev	Fr	Mean	StDev	Fr	Mean	StDev
0.15	3.632 8	2.16 %	0.24	5.8403	1.07 %	0.34	11.613 8	0.98 %
0.19	4.287 7	1.33 %	0.29	8.3442	1.92 %	0.39	16.987 6	1.21 %

The form factor k is calculated as recommended by the ITTC (1978) using Prohaska's method (ITTC, 1966). The approach assumes that CR can be approximated by the following equation at low speed and without separation:

$$C_W = C_{TM} - C_V = C_{TM} - C_{FM}(1+K) \quad (2)$$

$$\frac{C_T}{C_F} = (1+K) + \frac{yFr^4}{C_F} \quad (3)$$

and values of C_T/C_F form a line with slope y and intercept $(1+k)$. For The hull model, C_T data is used for $0.05 \leq Fr \leq 0.2$, and the results are provided below in figure 17. The estimated intercept is $(1+k)=1.207$.

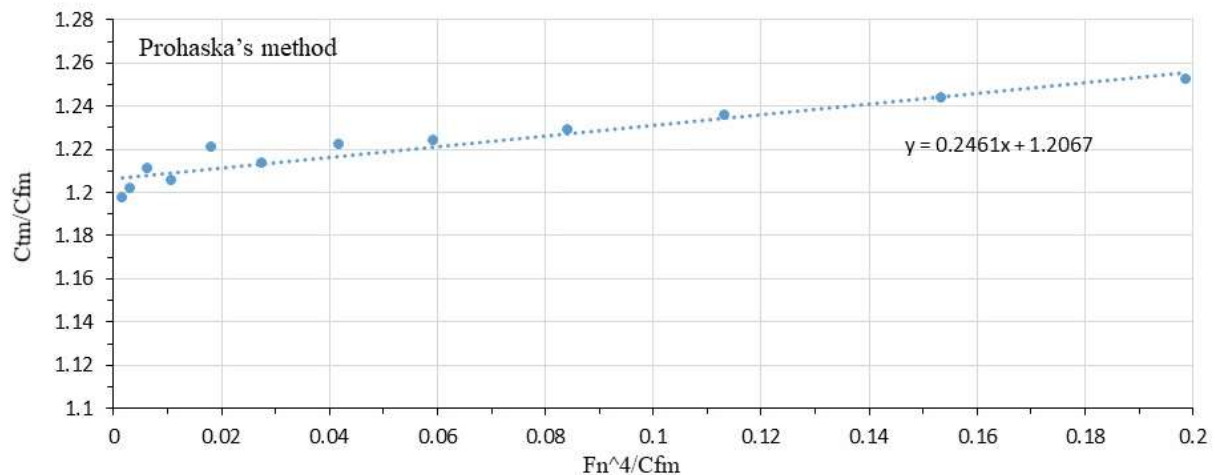


Figure 31 Evaluation of The Form Factor for Model 5415-51 Using Prohaska's Method

The graphs of total and residual resistance coefficients (C_{TM} and C_R vs. F_r , respectively, are shown in figure 18. Also viscous resistance has reported according to the ITTC 57 formula. Even though the model was unloaded and demounted at the end of each day, the data obtained reveal no appreciable dispersion. When F_r is between 0.05 and 0.15, C_{TM} gradually drops, remains constant ($C_{TM}=0.005$) between 0.15 and 0.225, gradually increases between 0.225 and 0.35, and rapidly increases when F_r is greater than 0.35. For F_r values of 0.1 to 0.225, 0.225 to 0.35, and F_r greater than 0.35, C_R is essentially piecewise linear with a rising slope. The presence of humps and hollows is minimal in both situations.

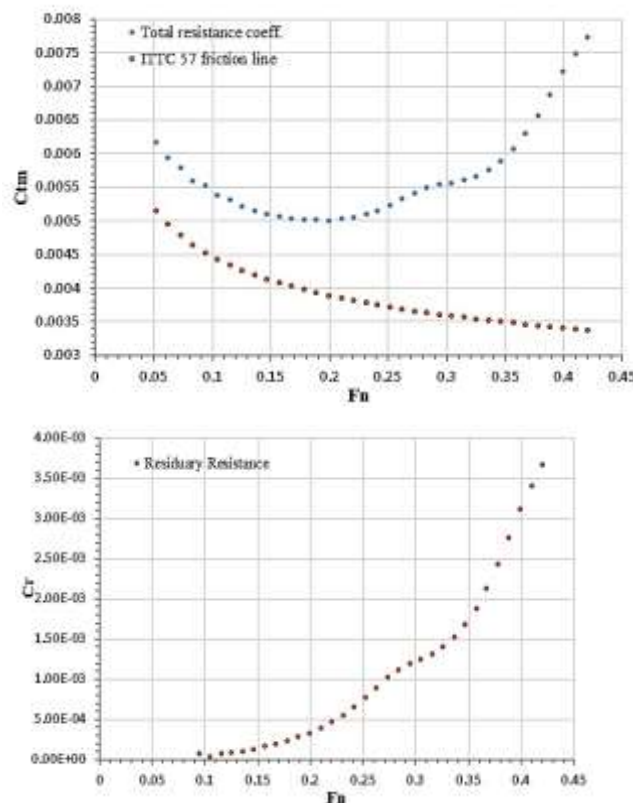


Figure 32 The Curves of Total and Residual Resistance Coefficient C_{TM} and C_R

As a function of carriage speed, a series of towing tests were conducted between $Fr = 0.05$ and 0.40 . The model was allowed to take its natural sinkage and trim angle during the tests. The average of 10 times-repeated experiments yielded the values at $Fr = 0.10, 0.20$, and 0.30 . Sinkage and model resistance have both been measured fore and aft. The sinkage σ and trim τ versus Fr are shown in figures 19 a & b. Model sinkage rises as Fr increases, and she has a weak bow down trim for Fr less than 0.20 and a strong bow up trim for Fr greater than 0.20 . σ is roughly linear for $0.1 < Fr < 0.35$. For $Fr < 0.2$, the data is mildly increases, whereas from $Fr 0.2$ to 0.35 , σ rapidly increases. τ is roughly piecewise linear with mild negative slope for $0.1 < Fr < 0.26$ and increasing positive slope for $0.26 < Fr < 0.4$.

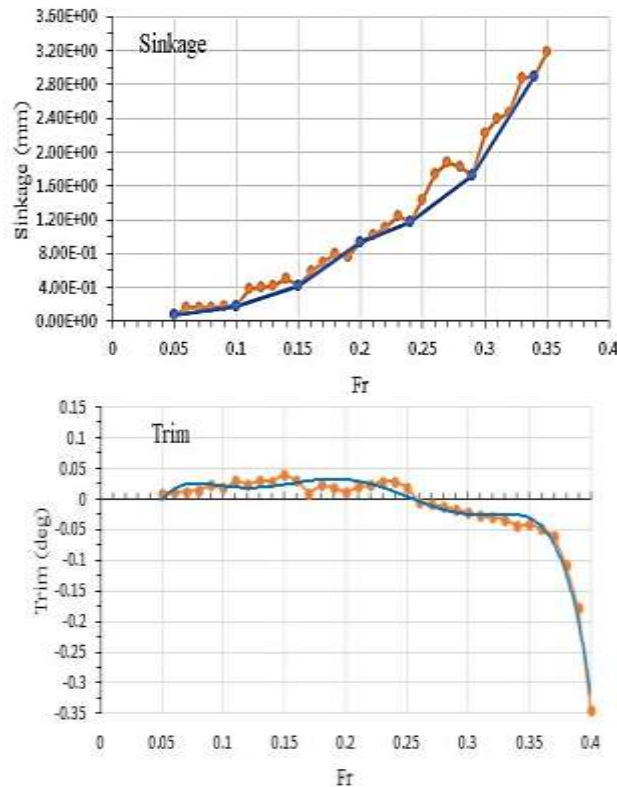


Figure 33 Sinkage And Trim Data for Model 5415-51

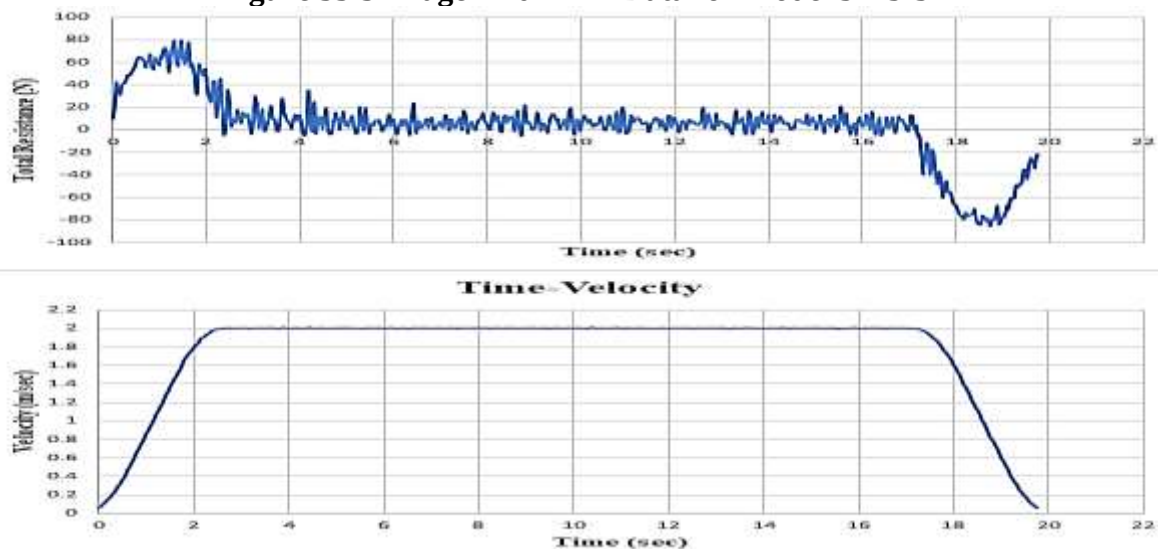


Figure 34 Time Resistance Chart – Time Velocity Chart At 2 M/Sec

Figure 20 shows the time resistance chart and the time velocity chart whereas the measured speed and drag vs time at 2 m/sec in MTC-TT are demonstrated in figure 21.

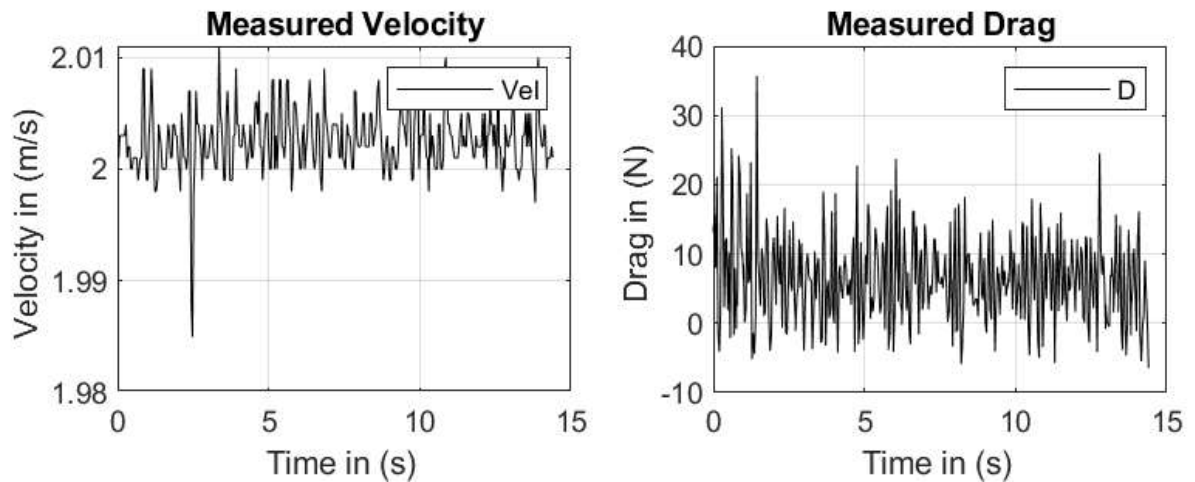


Figure 35 The Measured Speed and Drag Vs Time At 2 M/Sec In MTC-TT

The wave pattern and the wave outlines along the hull at $Fr = 0.30$ are shown in figures 22 and 23 based on data from the towing tank. When it comes to accurately capturing the general shape of the wave pattern, we observe that the towing tank tests are in good agreement with the CFD code.

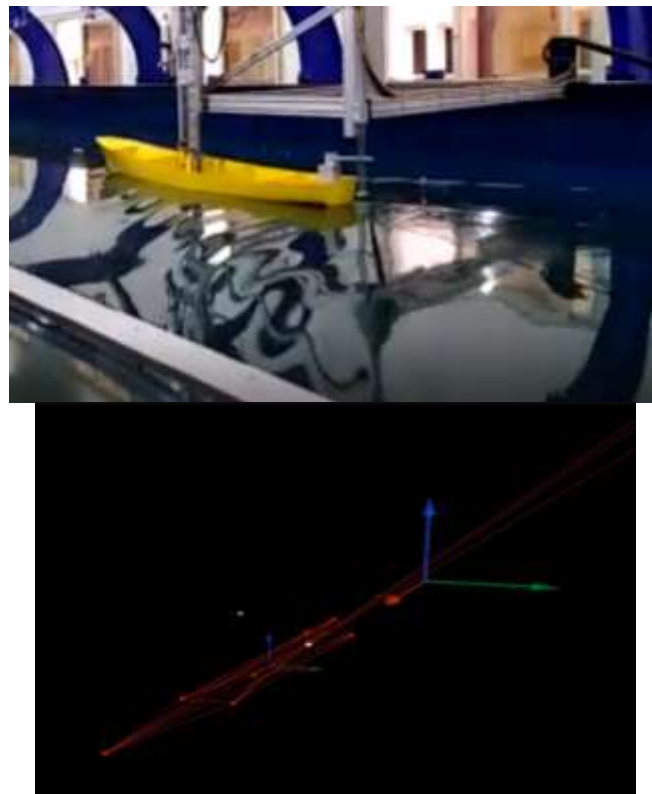


Figure 36 The Wave Pattern Along the Hull At $Fr = 0.30$



Figure 37 The Wave Contours Along the Hull at $Fr = 0.30$

Figure 24 shows the CFD results as well as the experimental data from the towing tank for total resistance. As a way of predicting calm water resistance, it offers a graphic representation of the agreement between experimental and CFD values. Table 6 compares the experimental results to the numerical estimates of ship resistance, and the error % is provided by

$$\Delta R_T\% = (R_T^{CFD} - R_T^{exp}) / R_T^{exp}. \quad (5)$$

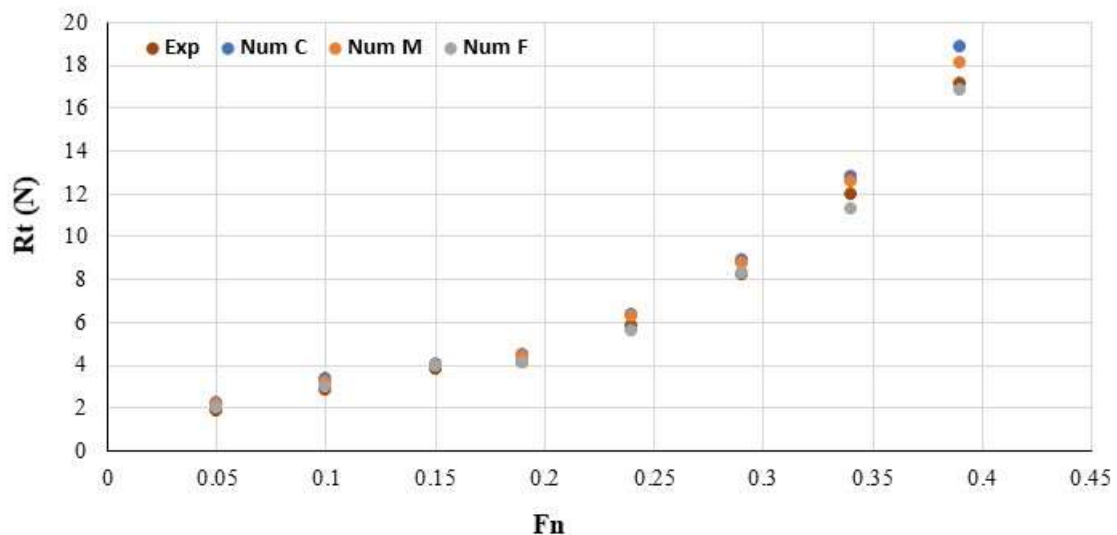


Figure 38 Total Resistance Experimental Data Plotted with CFD Results

At lower speeds, the agreement of the calculations from the CFD simulations with the experimental data worsens, and as shown in Table 6, at $Fr=0.35$, the CFD values deviate significantly from the experimental data. In general, the results show that the resistance values at various speeds for the fine mesh show a very good and improved agreement with the

experimental data, with an error of less than 6%. The highest agreement between experimental and computational results is found for total resistance.

Table 11 The Comparison of Experimental and CFD Results for DTMB #5415-51 Using Different Meshes

F_n	V_m (m/s)	R_T (N) (Exp.)	R_T (N) (C) (Simulation)	ΔR_T (%)	R_T (N)(M) (Software)	ΔR_T (%)	R_T (N) (F) (Software)	ΔR_T (%)
0.05	0.25	1.89269414	2.263521	19.59254	2.152487	13.72609	1.995422	5.4275785
0.1	0.5	2.88282046	3.391023	17.62866	3.254165	12.88129	2.987445	3.6292421
0.15	0.75	3.78658455	4.008125	5.850667	4.00052	5.649826	3.985412	5.2508391
0.19	1	4.153801	4.514248	8.677522	4.45213	7.182073	4.105481	-1.163272
0.24	1.25	5.82283543	6.392025	9.775127	6.25418	7.407809	5.652481	-2.925627
0.29	1.5	8.24017278	8.921542	8.268871	8.744121	6.115748	8.312584	0.8787585
0.34	1.75	11.9915089	12.77402	6.525552	12.54152	4.586672	11.33254	-5.4952956
0.39	2	17.1598606	18.84354	9.811737	18.12484	5.623469	16.84521	-1.8336429

5. Conclusion

The objective of the current study is to contribute to combatant geometry DTMB Model 5415-51's surface-ship resistance, sinkage, and trim. Reliable CFD data is needed for a benchmark test conducted internally. It confirms the significance of such arduous experiments in understanding the behavior of the flow around the DTMB model. The primary goal of this experiment is to create a useful database for numerical code testing. The model's CFD codes have to undergo a thorough examination in order to verify the experimental results.

A significant portion of the current data set was cross-validated using experimental results obtained at DTMB, yielding a very good result. Extensive comparisons between our CFD calculations and experimental results for DTMB 5415-51 show that the method used is acceptable and that the mesh generation method can be used to estimate the model's hydrodynamic performance. The fine grid will be utilized since it is most appropriate for the calculation and produces the best results because convergent results are achieved as mesh size decreases, according to a comparison of total resistance grid results by CFD. On the other hand the obtained experimental results validated the setup of the towing tank facility and adequacy of procedures.

Statement and Declaration

The authors have no relevant financial or nonfinancial interests to disclose. The authors have no competing interests to declare that are relevant to the content of this article

References

- [1] Z. rong ZHANG, H. LIU, S. ping ZHU, and F. ZHAO, "Application of CFD in ship engineering design practice and ship hydrodynamics," J. Hydrodyn., vol. 18, no. 3 SUPPL., pp. 315–322, Jul. 2006.
- [2] F. Stern, J. Longo, R. Penna, A. Olivieri, T. Ratcliffe, and H. Coleman, "International collaboration on benchmark cfd validation data for surface combatant dtmb model 5415*," in Twenty-Third Symposium on Naval Hydrodynamics, 2001.
- [3] A. Shi, M. Wu, B. Yang, X. Wang, and Z. Wang, "Resistance calculation and motions simulation for free surface ship based on CFD," Procedia Eng., vol. 31, pp. 68–74, 2012.
- [4] K. Żelazny, "A method of calculation of ship resistance on calm water useful at preliminary stages of ship design," Zesz. Nauk., vol. 38, no. 110, pp. 125–130, 2014.
- [5] A. Elhadad, W. Y. Duan, R. Deng, and H. Elhanfey, "Numerical analysis for resistance calculations of NPL as a floating hull for wave glider," Appl. Mech. Mater., vol. 619, pp. 38–43, 2014.
- [6] F. Zhao, S. P. Zhu, and Z. R. Zhang, "Numerical experiments of a benchmark hull based on a turbulent free-surface flow model," C. - Comput. Model. Eng. Sci., vol. 9, no. 3, pp. 273–285, 2005.
- [7] ITTC, "Resistance Committee 1.," Resist. Comm. 27th ITTC, no. February 2012, pp. 27–28, 2014.
- [8] W. Rosemurgy, D. Edmund, K. Maki, and R. Beck, "A Method for Resistance Prediction in the Design Environment," in 11th International Conference on Fast Sea Transportation, FAST 2011 - Proceedings, 2011.
- [9] Y. Ahmed and C. G. Soares, "Simulation of the Flow around the Surface Combatant DTMB Model 5415 at Different Speeds," 13th Congr. Int. Marit. Assoc. Mediterr. (IMAM 2009), no. January, 2009.
- [10] F. R. Menter, "Two-equation eddy-viscosity turbulence models for engineering applications," AIAA J., vol. 32, no. 8, pp. 1598–1605, 1994.
- [11] A. Elhadad, W. Y. Duan, and R. Deng, "Comparative Investigation of an Automated Oceanic Wave Surface Glider Robot Influence on Resistance Prediction Using CFD Method," Appl. Mech. Mater., vol. 710, pp. 91–97, Jan. 2015.
- [12] A. Elhadad, "Resistance Evaluation for the Submerged Glider System using CFD Modelling," J. Adv. Res. Appl. Sci. Eng. Technol., vol. 3, no. 3, pp. 147–159, 2023.
- [13] S. H. Rhee, "COMPUTATIONAL VALIDATION OF FLOW AROUND SURFACE SHIPS USING," 2016, no. January 2006.
- [14] M. Diez et al., "Multi-objective Hydrodynamic Optimization of the DTMB 5415 for Resistance and Seakeeping," 2020, pp. 0–17.
- [15] L. Larsson, F. Stern, and M. Visonneau, "CFD in ship hydrodynamics - Results of the Gothenburg 2010 workshop," in Computational Methods in Applied Sciences, 2013, vol. 29, pp. 237–259.
- [16] D. A. Jones and D. B. Clarke, "Fluent Code Simulation of Flow around a Naval Hull: the DTMB 5415," Dsto-Tr-2465, 2010.
- [17] F. Pacuraru, "Validation of potential flow method for ship resistance prediction," in IOP Conference Series: Materials Science and Engineering, 2019, vol. 591, no. 1.
- [18] E. Methods, H. Speed, M. Vehicles, and S. K. Tests, "ITTC - Recommended Procedures and Guidelines ITTC - Recommended Procedures and Guidelines High Speed Marine

- Vehicles : Sea Keeping Tests,” ReVision, 2005.
- [19] A. Olivieri, F. Pistani, A. Avanzini, F. Stern, and R. Penna, “Towing Tank Experiments of Resistance , Sinkage and Trim , Boundary Layer , Wake , and Free Surface Flow Around a Naval Combatant Insean 2340 Model,” Security, no. 421, pp. 1–42, 2001.
 - [20] E. Begovic, a. H. Day, and A. Incecik, “Experimental Ship Motion and Load Measurements in Head and Beam Seas,” in Proceeding of The 9th Symposium on High Speed Marine Vehicles, 2011, pp. 1–8.
 - [21] T. Pedersen, S. Nylund, and A. Dolle, “Wave height measurements using acoustic surface tracking,” in Oceans Conference Record (IEEE), 2002, vol. 3, pp. 1747–1754 vol.3.
 - [22] ITTC, “Ship Models – 7.5-01-01-01,” 23th Int. Towing Tank Conf., pp. 266–273, 2002.
 - [23] Y. Toda, F. Stern, I. Tanaka, and V. Patel, “Mean-Flow Measurements in the Boundary Layer and Wake of a Series 60 Block Coefficient = .6 Model Ship with and without Propeller,” J. Sh. Res., vol. 34, p. 109, 1988.
 - [24] J. Andersson, M. Hyensjö, A. Eslamdoost, and R. E. Bensow, “CFD Simulations of the Japan Bulk Carrier Test Case,” 2015.
 - [25] K. Suastika, “Comparative Resistance Test Between Two Towing Tanks (a Case Study At Its and Ih1),” 2018, no. January, pp. 308–317.
 - [26] D. Purnamasari, I. K. A. P. Utama, and K. Suastika, “Benchmark Study of Ship Model Resistance Test,” Appl. Mech. Mater., vol. 874, no. January, pp. 114–120, 2018.
 - [27] Y. Ozdemir, B. Barlas, T. Yilmaz, and S. Bayraktar, “Numerical and experimental study of turbulent free surface flow for a fast ship model,” Brodogradnja, vol. 65, pp. 39–54, 2014.
 - [28] ITTC, “T h e R e s i s t a n c e C o m m i t t e e,” Proc. 28th ITTC, 2017.
 - [29] ITTC, “Fresh Water and Seawater Properties - 7.5-02-02-01.02,” 26th Int. Towing Tank Conf. Rio Jenerio, Brazil, 28 August - 3 Sept., pp. 1–45, 2011.
 - [30] E. Date, “ITTC – Recommended Procedures and Guidelines ITTC Quality System Manual Recommended Procedures and Guidelines Example for Uncertainty Analysis of Resistance Tests in Tow- ing Tanks ITTC – Recommended Procedures and Guidelines,” 2014.
 - [31] T. Gourlay, J. H. Ha, P. Mucha, and K. Uliczka, “Sinkage and Trim of Modern Container Ships in Shallow Water,” 2015.
 - [32] J. Sun, H. Tu, Y. Chen, D. Xie, and J. Zhou, “A Study on Trim Optimization for a Container Ship Based on Effects due to Resistance,” J. Sh. Res., vol. 60, pp. 30–47, 2016.
 - [33] S. Sherbaz and W. Duan, “Ship Trim Optimization: Assessment of Influence of Trim on Resistance of MOERI Container Ship,” Sci. World J., vol. 2014, p. 603695, 2014.
 - [34] P. Akbarzadeh, P. Molana, and M. A. Badri, “Determining resistance coefficient for series 60 vessels using numerical and experimental modelling,” Ships Offshore Struct., vol. 11, no. 8, pp. 874–879, 2016.
 - [35] S. Mancini, C. Pensa, F. De Luca, and S. Miranda, “An Extended Verification and Validation Study of CFD Simulations for Planing Hulls,” J. Sh. Res., vol. 60, pp. 101–118, 2016.

DESIGN OF WATER QUALITY MONITORING SYSTEM BASED ON INTERNET OF THINGS TECHNOLOGY

Norsuzila Ya'acob^{1,2,*}
Nur Syaza Zainali¹
Amirul Asraf Abdul Rahman¹
Azita Laily Yusof^{1,2,3}
Murizah Kassim^{1,3}
Ahmad Shazri Nazif Salehudin¹

¹School of Electrical Engineering, College of Engineering, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia

²Wireless Communication Technology (WiCoT), School of Electrical Engineering, College of Engineering, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia

³Institute for Big Data Analytic and Artificial Intelligence (IBDAAI), University Teknologi MARA, 40450 Shah Alam, Malaysia

Abstract: Water quality is an assessment of how appropriate water is for a certain use or purpose, taking into consideration various physical, chemical, and biological factors that can affect its suitability. These factors can include pH, turbidity, dissolved oxygen, temperature, and the presence of pollutants or pathogens. The outdated method has been used by scientists and researchers to monitor the quality of water from the sources. The objective of this project is to create an efficient Internet of Things (IoT) system that can various sensors to continuously monitor water quality. The system is implemented using Arduino as the microcontroller, and sensors. A real-time monitoring system that is IoT-based was done to improve the examination process of the water sample. The system device is containing a NodeMCU ESP8266 microcontroller, pH, temperature, and turbidity sensors and uses the Blynk application. The system experiment results show that the device can show different readings based on the variety of water samples from different water bodies.

Keywords: Water Quality Monitoring System, Iot, Ph, Temperature, Turbidity, Arduino, ESP8266, Blynk

1. Introduction

Water is the source of life; it has a deep impact on human health. About 70% of Earth's surface is covered with water, yet the amount of fresh water fit for human consumption and usage is as low as 2% of the total volume [1-2]. As society and the economy have rapidly advanced, humans are polluting the water with too much sewage, and it's causing problems for the environment [3-4]. Water quality can change based on location, time, weather, and pollution sources, making it vital to frequently monitor and test the water to ensure it meets the necessary standards for its intended use. The work of scientists, researchers, and water resource managers are crucial in monitoring, preserving, and enhancing water quality [5-6]. Water quality can also be evaluated for the presence of contaminants such as pesticides, herbicides, heavy metals, and algae [7]. Monitoring water quality helps to identify the impacts of human activity on the environment and to understand and predict natural processes. It can also be used to support restoration efforts and ensure that environmental regulations are being followed.

In the past, the water quality of an environmental system can be determined through a manual process, where samples of water are collected using tools such as glassware, scoops, and bottles. The samples are then analyzed in a laboratory to measure various parameters and determine the water content. This process can take anywhere from 1 to 30 days, depending on the specific measurements being made [8]. Chemical indicators are one of the methods were used to determine the acidity of a certain liquid back in the day. These indicators will change colour to indicate whether the water or liquid is acidic or alkaline [9]. It is too hard to decide the pH level accurately because of the limitations of the human eyes. The same goes for temperature where the usage of a physical thermometer has a higher chance of parallax error. The invention of electronic sensors has helped humans, especially researchers to measure and determine the levels of water quality accurately with precise calculations.

The fast development of the Internet of Things has created many opportunities and solutions to many problems, one of which is water quality monitoring. IoT technology is well suited for monitoring water quality and the environment, due to its advanced intelligence, real-time data capabilities, wide range of coverage, ability to simultaneously gather data from multiple sensors, and scalability. It has a wide range of potential applications in this field [10].

Sui H. et al. [11] used the application of NB-IoT technology to monitor the water quality in the city's open water by collecting data from several different points using multiple devices. Shanmugam K. *et al.* [12] are using the Google Firebase application as a platform for their smart water quality monitoring to make the application more user-friendly and interactive. Hakimi I. et al. [13] are using Arduino Uno and Blynk applications to develop their monitoring device at low-cost. Sung W. *et al.* [14] develop a water quality monitoring system using physiochemical sensors to get more accurate measurements. Prabowo O. *et al.* [15] designed a remote-operated vehicle system to monitor water quality.

This project aims to design a system with modern features of water quality checking and water usage monitoring systems. The parameters that will be measured are pH, turbidity, and temperature and the relationship between them using sensors that are designated to each parameter. All the input will be processed using the Arduino assisted by a Wi-Fi module of NodeMCU to be controlled through a wireless connection from the Blynk apps on the smartphone.

2. Methodology

2.1 Framework

This section will explain the process of how the water quality monitoring system works as shown in Figure 1. The Arduino board, which is based on the ATmega328P microprocessor, serves as the system's primary processor. The characteristics of water quality are gathered using a turbidity sensor, a pH sensor, and a temperature sensor. After being processed by Arduino, the collected data is transmitted to a Blynk application in the smartphone through the ESP8266 Wi-Fi module for real-time monitoring.

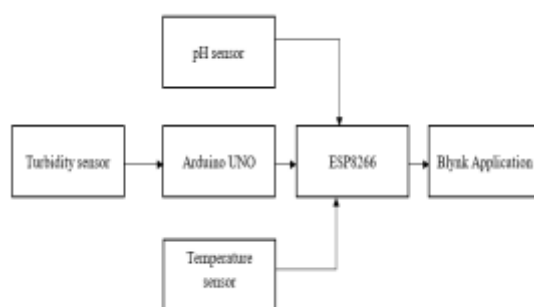


Figure 39: System's Framework Flowchart

2.2 Materials

The water samples were gathered from various places such as tap water, ditch, soap water, lime juice. The parameters that will be measured are pH, temperature, and turbidity. Further testing will be done on the ditch water at the address Lot 1733B, Batu 7, Kampung Sijangkang at coordinate 2.94825738971433, 101.46138037883536 to monitor the water quality. The water samples are prepared inside plastic bowls and then the samples will be put outside the house to be exposed in the same condition. Apply to all water samples except for ditch water where it will be gathered every period shown in the result section. The data for all the water samples will be compared to the standard of water quality based on The Engineering Services Department of the Ministry of Health of Malaysia. The standard pH and turbidity that are recommended drinkable are in the range of 6.5 – 9.0 for pH and 5 NTU or less for turbidity [16-17], [18-19]. The temperature of the water was measured because the water quality also can be affected, especially the aquatic life, by the rate of microbiology growth that depends on the water temperature [20-21]. Table 1 shows the standard range of the sensors.

Table 1: Standard range of the sensors

Parameter	Standard range
pH	6.5 – 9
Temperature	24 – 32°C
Turbidity	0 – 2600 NTU

2.3 Software

The monitoring app on a smartphone serves as the software element of the suggested system. This section describes how to connect to an Esp8266 through Wi-Fi using Blynk applications. Additionally, temperature, pH, and turbidity sensor data can be read by the Esp8266. The flowchart of the system software is shown in Figure 4. The user must first connect to Wi-Fi and the Blynk app on their smartphone to begin the software portion. The Esp8266 component must be interfaced with the Blynk app to connect to Wi-Fi. After reading data from sensors connected to the microcontrollers, the data will then be shown on the user's smartphone.

2.3.1 Arduino IDE

The Arduino Integrated Development Environment (IDE), shown in Figure 2, also known as the Arduino Software (IDE), includes a code editor, a message area, a text console, a toolbar with buttons for common functions, and a series of menus. It communicates with and uploads programs to the Arduino hardware [22].



Figure 2: Arduino IDE

2.3.1 Blynk Application

Blynk (Figure 3) is an IoT platform that allows users to control devices such as Arduino and NodeMCU via a smartphone app on iOS or Android. It allows users to create a graphical user interface by using widgets to display sensor data and remotely control hardware, as well as store and visualize data. It is specifically designed for the Internet of Things [23].



Figure 3: Blynk Application

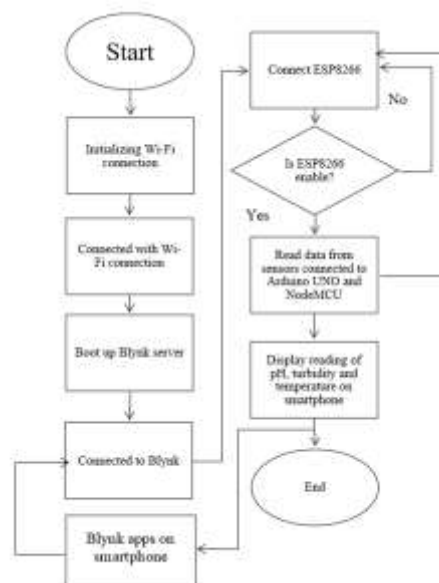


Figure 4: The Software Flowchart

2.4 Hardware

For the hardware part, the suggested system required an Arduino UNO R3 controller, an ESP8266 Wi-Fi module, a pH sensor, a turbidity sensor, and a water temperature sensor. Figure 5 illustrates the functioning of the suggested system, which is a water quality monitoring system that is based on software and hardware. The process starts with the initialization of the three sensors, pH, turbidity, and temperature. Then, the sensors will function to read and transfer data through the microcontroller Arduino and send the data to the Blynk application through the Wi-Fi module to be displayed. The equipment includes an Arduino Uno, Esp8266 Wi-Fi module, and smartphone which enable us to continuously determine the characteristics of water, as well as pH and turbidity sensors. The monitoring app on a smartphone serves as the software element of the suggested system.

Readings are straightforward and consistent because every single sensor is submerged in water for testing. We need to use an Analog Digital Converter (ADC), which is already present in the AT Mega 328P in the Arduino Uno, to convert the qualities from simple to advanced and discrete because they are simple. The characteristics will be displayed in the Blynk app, and the Esp8266 will provide microcontroller access to Wi-Fi. Figure 5 shows the flowchart of the hardware.

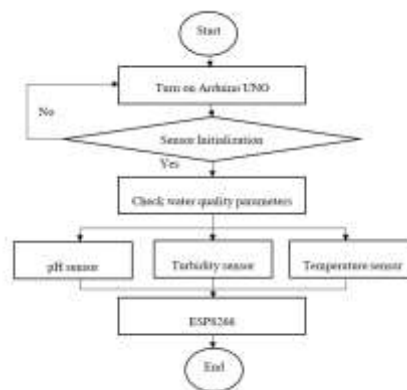


Figure 5: The Hardware Flowchart

2.5 Circuit

Figure 6 below shows the circuit diagram of the water quality monitoring system. To explain the diagram, the pH sensor module is connected to the analogue pin in the NodeMCU and the temperature sensor is connected to the digital pin. Both sensors are powered by the NodeMCU. Meanwhile, the turbidity sensor module is connected to the analogue pin in the Arduino UNO and digital pins 2 and 3 are connected to the Rx/Tx pin of the NodeMCU to be able to connect to the Wi-Fi.

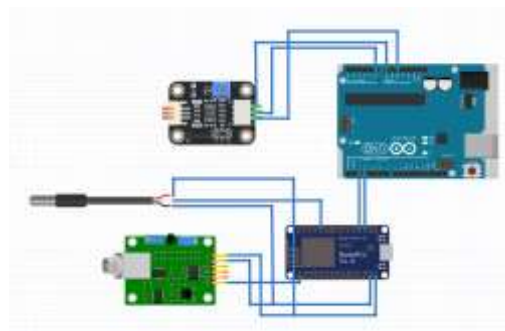


Figure 6: The Circuit Diagram

3. Results

3.1 Hardware

Figure 7 shows the condition of the hardware when it was running. The LED light is the indicator that the hardware is functioning as intended.

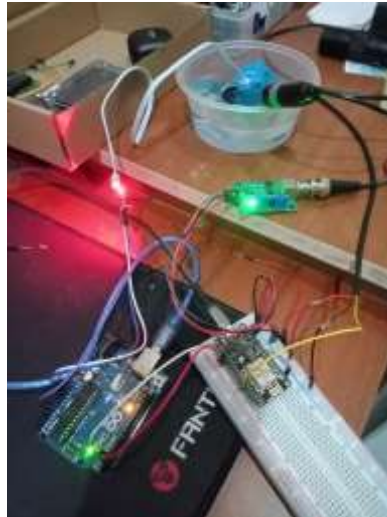


Figure 6: The Hardware Result

3.2 Software

The test of the system was done by taking multiple water samples from various water bodies to ensure it is properly functioning. The data was collected on January 30, 2023, at 8.00 am. The reading of the measurements was displayed on the label widget in the Blynk website. To test the accuracy of the sensors, the sensors are set to take 800 samples at one time and then calculate the average value which will be recorded. The then also be 3 times to get the average value. The result of the system is shown in Table 2 while Table 3 is the average value.

Table 2: PH Turbidity temperature result

Water sample	pH	Turbidity (NTU)	Temperature (°C)
Tap water	7.19	162.31	26.32
	7.20	156.16	26.18
	7.25	150.04	26.20
Ditch	8.14	1724.54	25.89
	8.25	1756.23	26.05
	8.19	1802.84	26.14
Lime juice	4.56	725.25	26.43
	4.66	731.14	26.48
	4.53	733.26	26.52
Dish soap water	8.02	424.11	26.20
	8.09	437.36	26.23
	8.05	436.43	26.18

Table 3: PH Turbidity temperature result

Water sample	pH	Turbidity	Temperature (°C)
Tap water	7.21	16.82	26.23
Ditch	8.27	820.82	26.02
Lime juice	4.5	430.12	26.48
Dish soap water	8.05	359.61	26.2

Figure 8 shows the pH value of the 4 different water types that was tested. The test was done to check whether the sensor able to detect different pH value of different substances. Tap water has the pH value range between 7.19 – 7.25 with average of 7.21 which is ideal for domestic usage. The pH of lime juice is at the lowest pH value with the average of 4.5 as lime is containing with citric acid which is acidic. The soap water and ditch water have pH value above 7 with the soap water has average value of 8.05 because soap is an alkaline and ditch water 8.27.

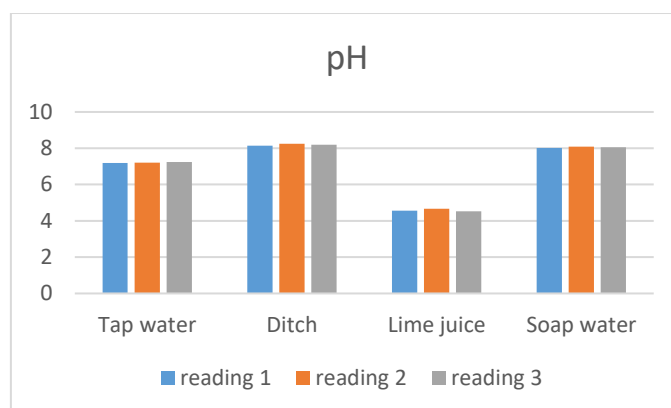


Figure 8: PH Of 4 Different Type of Water

Figure 9 shows the temperature of the 4 different types of water. Based on the observation, each water type has a quite inconsistent trend in the graphs. The change in temperature is because of thermodynamic equilibrium. When the water samples are put in the plastic bowls, and the temperature sensor merges into the water, heat transfer between the medium and surrounding happens [24]. At the same time, when mixing lime juice and dish soap into the water, its exothermic reactions happen and cause the temperature to increase.

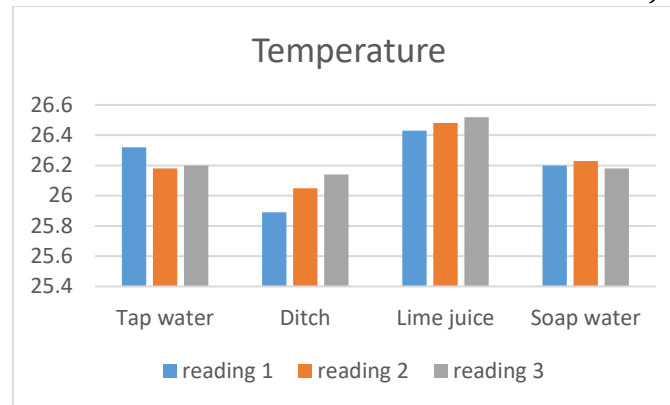


Figure 9: Temperature

Figure 10 shows the turbidity of water for all 4 types of water. The turbidity sensor is shown to be able to detect the murkiness of water. Ditch water has the highest turbidity with average value of 820.82 NTU. This is because the ditch water contains various kinds of substances in it including mud, algae and other microorganisms. Meanwhile tap water has lowest turbidity value with average of 16.82 NTU as the water is crystal clear and safe to for domestic usage.

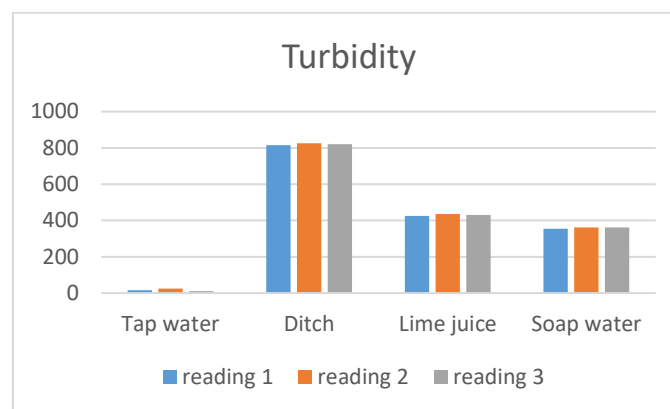


Figure 10: Turbidity

Figure 11 shows the results of the label widget in the Blynk website during the testing of the system. The NodeMCU microcontroller is to connect the system to the internet to monitor the measurement of the water samples using the Blynk website.



Figure 11: Blynk Website During Testing

Table 4, Table 5, and Table 6 shows the data obtained from ditch water on January 30, 2023 for 3 different period monitoring processes at 8.30 am, 1.30pm and 7.30pm. This section will discuss the result gained from the testing.

Table 4: pH values gained for 3 different period monitoring process

Water sample	8.30 am	1.30 pm	7.30 pm
Tap water	7.19	6.97	7.09
Ditch	8.23	7.41	8.25
Lime juice	4.49	4.72	4.66
Dish soap water	8.01	7.85	8.11

Table 5: Temperature values gained for 3 different period monitoring process

Water sample	8.30 am	1.30 pm	7.30 pm
Tap water	27.25	29.38	28.53
Ditch	25.56	30.91	27.81
Lime juice	26.63	29.42	28.56
Dish soap water	26.69	29.35	28.31

Table 6: Turbidity values gained for 3 different period monitoring process

Water sample	8.30 am	1.30 pm	7.30 pm
Tap water	16.21	20.56	18.27
Ditch	816.54	1288.35	1089.26
Lime juice	721.98	913.69	822.42
Dish soap water	424.74	591.28	468.63

The results from Table 4, Table 5, and Table 6 were plotted on graphs as shown in Figure 12, Figure 13, and Figure 14 respectively. Figure 12 shows the change in pH value on that day for 3 different periods. The pH value throughout the experiment lies between over 7 to close to 9 except for lime juice. All the samples at its respective lowest pH at 1.30 pm respectively. The change in pH value is caused by the concentration of hydrogen ions present in the water[19]. When the concentration of hydrogen ions is high, the pH value decreases and vice versa. During midday, the sun was at its peak position to the Earth which caused temperatures to increase. This caused the molecular vibration of the water increase which results in the ability of water to ionized and form more hydrogen ions [25].

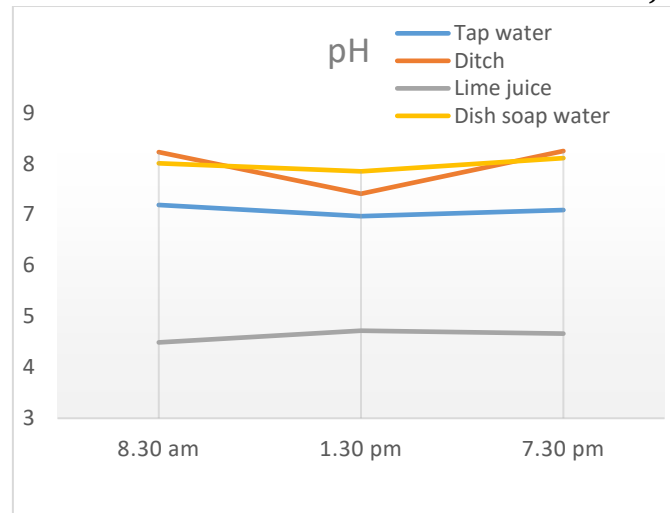


Figure 12: pH Graph

Figure 13 shows the graph of the temperature of the water during monitoring process. From the result, it shows that the trend of the temperature of the samples were similar where first the temperature increase at the beginning and decrease at the end. The highest temperature recorded is 30.91°C at 1.30 pm and lowest temperature is 25.56°C at 8.30 am from the water ditch sample. On that day, the sun was at its peak on midday at 1.30pm. On the afternoon, it starts getting cloudy from 3.30 pm onwards and cause the temperature to drop drastically until 7.30pm.

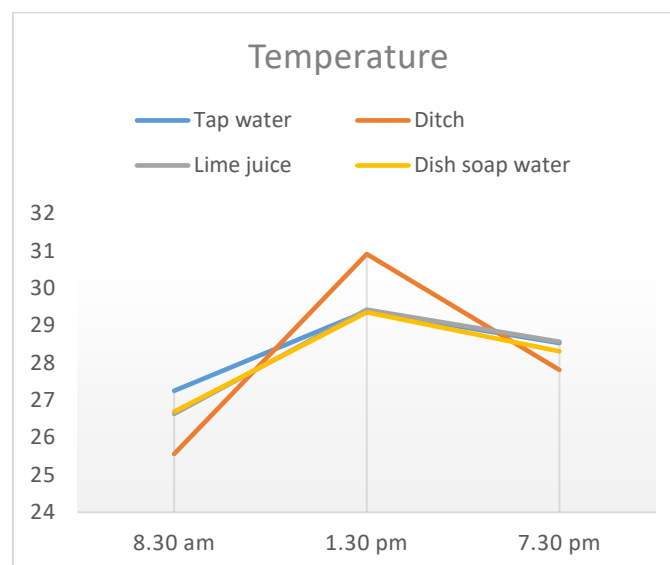


Figure 13: Temperature Graph

Figure 14 shows the graph of turbidity of the water changes throughout the day. The change of turbidity may cause by water flow in the ditch where it can drag away or bring in residue like mud or other solids substances to the testing spot. It also can be affected by human activities like deforestation. The graph shows that the turbidity is at the highest during midday when the temperature is also the highest. This condition tells that the temperature also effect the turbidity of the water. The reason for this to happen is the high temperature causing disruption to the sensors. Changes in temperature can cause variations in the refractive index of water, which can in turn affect the accuracy of optical sensors used to measure turbidity [26].

The other factors that temperature can give impact turbidity are the growth of algae and the decomposition of organic matter, which can increase the number of suspended particles [27] which may not be possible for this experiment since it was held for a short period of time.

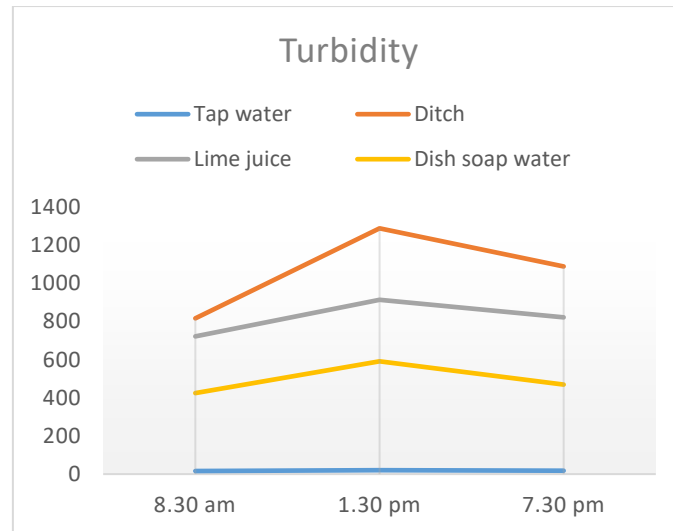


Figure 14: Turbidity9 Graph

4. Conclusion

The water quality monitoring system design that is based on the Internet of Things feature is low-cost, reasonable, rigorous, and functional. The hardware can connect to the internet and the system can show the readings of the parameters in real time from the Blynk website. Based on the results, it can be concluded that the temperature of the water is affecting the condition of the water in terms of pH value and turbidity. The temperature is inversely proportional to the pH value. As the temperature increases the pH value decrease. The temperature is directly proportional to the turbidity, and it stays true even though the turbidity sensor was affected by the temperature. A few improvements could be done to the system such as the design of the hardware where it can be easily carried around. Use even more advanced and high-sensitivity tolerance sensors to get more accurate results.

Acknowledgement

The author acknowledges the Universiti Teknologi MARA for funding under the Grant LESTARI (600-IRMI 5/3/LESTARI (035/2019). The authors also would like to thank the School of Electrical Engineering, College of Engineering, Universiti Teknologi MARA, Shah Alam, Selangor for their valuable support.

References

- [1] M. Kumar Jha, R. Kumari Sah, M. S. Rashmitha, R. Sinha, B. Sujatha, and K. v Suma, "Smart Water Monitoring System for Real-Time Water Quality and Usage Monitoring," in 2018 International Conference on Inventive Research in Computing Applications (ICIRCA), 2018, pp. 617–621. doi: 10.1109/ICIRCA.2018.8597179.
- [2] S. K. Vasudevan and B. Baskaran, "An improved real-time water quality monitoring embedded system with IoT on unmanned surface vehicle," *Ecol Inform*, vol. 65, Nov. 2021, doi: 10.1016/j.ecoinf.2021.101421.
- [3] C. Zhang, J. Wu, and J. Liu, "Water quality monitoring system based on Internet of Things," in *Proceedings - 2020 3rd International Conference on Electron Device and*

- Mechanical Engineering, ICEDME 2020, May 2020, pp. 727–730. doi: 10.1109/ICEDME50972.2020.00171.
- [4] Q. Bai, J. Wu, and C. Jin, “Theater Quality Online Monitoring System Based on Wireless Sensor Network,” in Proceedings - 2020 13th International Symposium on Computational Intelligence and Design, ISCID 2020, Dec. 2020, pp. 234–237. doi: 10.1109/ISCID51228.2020.00059.
- [5] S. Giri and Z. Qiu, “Understanding the relationship of land uses and water quality in Twenty First Century: A review,” J Environ Manage, vol. 173, pp. 41–48, 2016, doi: <https://doi.org/10.1016/j.jenvman.2016.02.029>.
- [6] M. Camara, N. R. Jamil, and A. F. bin Abdullah, “Impact of land uses on water quality in Malaysia: a review,” Ecol Process, vol. 8, no. 1, p. 10, 2019, doi: 10.1186/s13717-019-0164-x.
- [7] “A Primer on Water Quality,” 2001. [Online]. Available: <http://www.epa.gov/safewater>
- [8] Slamim, Universitas Negeri Jember, Institute of Electrical and Electronics Engineers. Indonesia Section, Institute of Electrical and Electronics Engineers. Indonesia Section. Computer Society Chapter, and Institute of Electrical and Electronics Engineers, Proceedings, 2019 International Conference on Computer Science, Information Technology, and Electrical Engineering (ICOMITEE 2019): October 16th-17th 2019, Jember, Indonesia.
- [9] “chemical indicator -- Britannica Online Encyclopedia,” Encyclopaedia Britannica, Inc, Apr. 06, 2021. <https://www.britannica.com/science/chemical-indicator> (accessed Jan. 10, 2023).
- [10] Y. Wang and F. Lv, “Design of water quality monitoring system based on NB-IoT technology,” in 2022 3rd International Conference on Computer Vision, Image and Deep Learning & International Conference on Computer Engineering and Applications (CVIDL & ICCEA), 2022, pp. 825–828. doi: 10.1109/CVIDLICCEA56201.2022.9824243.
- [11] H. Sui, G. Zheng, J. Zhou, H. Li, and Z. Gu, “Application of NB-IoT technology in city open water monitoring,” in Proceedings - 2020 6th International Symposium on System and Software Reliability, ISSSR 2020, Oct. 2020, pp. 95–98. doi: 10.1109/ISSSR51244.2020.00023.
- [12] K. Shanmugam, M. E. Rana, and R. S. J. Singh, “IoT-based Smart Water Quality Monitoring System for Malaysia,” in 2021 3rd International Sustainability and Resilience Conference: Climate Change, 2021, pp. 530–538. doi: 10.1109/IEEECONF53624.2021.9668120
- [13] I. M. Hakimi and Z. Jamil, “Development of Water Quality Monitoring Device Using Arduino UNO,” IOP Conf Ser Mater Sci Eng, vol. 1144, no. 1, p. 012064, May 2021, doi: 10.1088/1757-899x/1144/1/012064.
- [14] W. T. Sung and F. N. Fadillah, “Water quality monitoring using physio chemical sensors,” in Proceedings - 2020 International Symposium on Computer, Consumer and Control, IS3C 2020, Nov. 2020, pp. 440–443. doi: 10.1109/IS3C50286.2020.00119.
- [15] O. M. Prabowo, Y. Tresnawati, and D. Kusumastuti, “Design of Water Quality Monitoring System using Remote Operated Vehicle,” in 7th International Conference on ICT for Smart Society: AIIoT for Smart Society, ICISS 2020 - Proceeding, Nov. 2020. doi: 10.1109/ICISS50791.2020.9307588.
- [16] Afif Sani, “National Standard For Drinking Water Quality Malaysia,” SaniChem Resources, May 07, 2022.
- [17] S. Jones, “Turbidity,” H2Olabcheck, Mar. 04, 2020. <https://www.h2olabcheck.com/blog/view/turbidity> (accessed Jul. 05, 2022).

- [18] E. Cirino and N. Butler, "What pH Should My Drinking Water Be?," Healthline, Mar. 29, 2019. <https://www.healthline.com/health/ph-of-drinking-water> (accessed Jul. 05, 2022).
- [19] J. Johnson and N. Olsen, "The pH of water: What to know," Medical News Today, Dec. 02, 2019. <https://www.medicalnewstoday.com/articles/327185> (accessed Jul. 05, 2022).
- [20] A. Card, K. Fitch, D. Kelly, C. Kemker, and K. Rose, "Water Temperature," Fondriest Environmental, Inc., Feb. 07, 2014. <https://www.fondriest.com/environmental-measurements/parameters/water-quality/water-temperature/> (accessed Jul. 05, 2022).
- [21] L. Rosenberg, "Heat Waves Are Taking a Toll On PNW Drinking Water," Green Matters, Jul. 14, 2021. <https://www.greenmatters.com/p/how-heat-waves-affect-drinking-water#:~:text=Hot%20weather%20can%20contaminate%20plastic%20bottled%20water.&text=And%20according%20to%20National%20Geographic,left%20sitting%20in%20the%20sun.> (accessed Jul. 05, 2022).
- [22] Arduino, "Arduino Integrated Development Environment (IDE) v1," Arduino, Dec. 15, 2022.
- [23] E. Media's, . Syufrijal, and M. Rif'an, "Internet of Things (IoT): BLYNK Framework for Smart Home," KnE Social Sciences, vol. 3, no. 12, Mar. 2019, doi: 10.18502/kss.v3i12.4128.
- [24] P. P. Urone and R. Hinrichs, Physics. Houston, Texas: OpenStax, 2020. Accessed: Jan. 30, 2023. [Online]. Available: <https://openstax.org/books/physics/pages/12-1-zeroth-law-of-thermodynamics-thermal-equilibrium>
- [25] "How Does Temperature Affect pH?," Westlab, Nov. 29, 2016.
- [26] M. Shi, J. Ma, and K. Zhang, "The Impact of Water Temperature on In-Line Turbidity Detection," Water (Basel), vol. 14, no. 22, p. 3720, Nov. 2022, doi: 10.3390/w14223720.
- [27] Water Science School, "Temperature and Water," U.S. Geological Survey, Jun. 06, 2018. <https://www.usgs.gov/special-topics/water-science-school/science/temperature-and-water> (accessed Jan. 29, 2023).

A MULTI-CRITERIA ENODE-B PRE-SELECTION AND HANDOVER TRIGGERING SCHEME FOR LTE-A HETEROGENEOUS WIRELESS NETWORK

Nur Syaza Zainali¹
Azita Laily Yusof^{2*}
Muhammad Aiman Zainali³
Norsuzila Ya'acob⁴

¹School of Electrical Engineering, College of Engineering, Universiti Teknologi MARA, Shah Alam, Selangor, Malaysia

Abstract: To establish the best connection possible, different radio access technologies are combined to form a heterogeneous wireless network (HetNets). The HetNets consists of macro cells and small cells in the Long Term Evolution-Advanced (LTE-A) network. However, the fast pace of development of LTE-A technology has been very challenging due to the increase in the number of user equipment (UE). Managing the selection of the target cell and migration of active communication sessions of the UE between different base station (BS) points is vital for seamless mobility. To address this issue, a multi-criteria eNodeB pre-selection parameter is introduced to pre-select the most suitable cell for handover. The proposed approach aims to enable UE to autonomously assess a list of potential eNodeB located in their direction of movement and choose the most suitable candidate network. A framework for eNodeB selection is developed that considers the distance between the candidate network and the vehicle's trajectory, as well as the mobility information of the vehicle. The important parameters which are the proximity index and distance factor are chosen to identify the best candidate eNodeB. Moreover, we investigate an efficient scheme for the handover trigger to get an optimum distance for handover initiation using the particle swarm optimization (PSO) technique. The results show that the proposed parameters are reliable for eNodeB selection to self-select the network and improved the successful handover probability by using the proposed optimum distance to initiate the handover process.

Keywords: Handover trigger; multi-criteria pre-selection; proximity index; distance factor

1. Introduction

The demand for wireless services and applications has grown tremendously in recent years due to the need for accessing information at any time and from anywhere. The development of smartphone applications in 2007 and 2008 further accelerated this growth. As a result, the need for a system with greater data capacity has become inevitable [1]. These advancements in wireless networking and communication technologies have led to the development of various types of wireless communication systems. These include Bluetooth for personal areas, IEEE 802.11-based WLANs for local areas, Universal Mobile Telecommunications Systems (UMTS) for wider areas, and satellite networks for global networking. 5G networks will be characterized by the smooth coexistence of various access technologies, such as UMTS, HSPA, WiMAX, LTE, LTE-A and Wi-Fi, along with cells of varying sizes, such as macro, pico, and femtocells [2]. However, vehicles moving through such diverse networks may encounter several difficulties, such as frequent and unwanted handovers, selecting incorrect candidate

networks for handovers, and experiencing long, drawn-out handover procedures, which can lead to failed handovers [3]. Vehicular Ad hoc Networks (VANETs) have overcome traditional wired networks that offer better services [4]. In addition, there is revolutionary progress with the latest 5G technology that integrates multiple networks with different architecture layers [5][6][7][8][9][10].

The research community has been making significant progress toward the convergence of various radio access technologies, leading to several proposals addressing heterogeneous networks regarding handover management, procedures, and protocols. Moreover, since 2008, the Third Generation Partnership Project (3GPP) Group has introduced the Access Network Discovery and Selection Function (ANDSF) in its Release 8. This function enables a 3GPP-compliant UE to discover non-3GPP access networks, such as Wi-Fi and WiMAX. In subsequent 3GPP Releases (10 and 12), this functionality has advanced to the extent of supporting the UE to discover, select, and connect to both non-3GPP and 3GPP access networks, such as UTRAN, LTE, and HSPA.

In a 4G wireless environment, a mobile user can continue using their mobile device while moving from one point of attachment to another. This process is known as handover, which enables a UE to keep its connection active as it moves from the coverage of one network access point to another access point. In LTE, handovers are typically classified as hard handovers, which means that there is a brief interruption in service during the handover process. This interruption can occur during both intra and inter-handover. A novel user mobility model as discussed in N. Bahra *et al.*, [11] deploys a vector autoregression (VAR) model and gated recurrent unit (GRU) to predict the future trajectory of a user. This approach has the lowest prediction error compared to the existing method. The handover process can be broken down into three main steps: system discovery, handoff decision, and handoff execution. During the system discovery phase, a mobile terminal equipped with multiple interfaces needs to determine which networks are available and what services are offered in each network. During the handoff decision phase, the mobile device decides which network it should connect to. During the handoff execution phase, connections need to be seamlessly rerouted from the existing network to the new network. This requirement is based on the Always Best Connected (ABC) concept, which includes authentication, authorization, and transfer of the user's context information [12][13].

LTE-A smoothly evolved from LTE and boasts higher peak data rates (1Gbps in the downlink and 500Mbps in the uplink), and higher spectral efficiency (30bit/s/Hz in the downlink and 15bit/s/Hz in the uplink). It supports various deployment scenarios, enabling seamless coverage from macro cells to indoor areas. When selecting a network, especially in this LTE-A multi-tier heterogeneous environment, it is important to prioritize factors such as the direction of movement and the coverage area of available networks over other parameters like received power signal and available bandwidth. Therefore, this paper proposes several parameters for the pre-selection of the base station or called eNodeB in LTE-A heterogeneous wireless network before initiating the handover process. Once the target eNodeB has been selected, then the optimum distance will be calculated using the PSO technique for handover triggering. The remainder of this paper is organized as follows. Section 2 explains the related works followed by methodology in the next section. Results and discussion are explained in section 4. Finally, the conclusion is summarized in section 4.

2. Related Works

The selection of an appropriate underlying network technology is critical for maintaining a satisfactory QoS for ongoing communication, especially in a heterogeneous multi-tier environment such as V2I communication. Various parameters need to be considered while selecting an optimal network technology, including the received signal strength indicator (RSSI), the speed and movement direction of the vehicle, network bandwidth and traffic load, delay, usage cost, and more. RSSI is the most commonly used parameter for selecting a suitable network for V2I communication. The direction and speed of the vehicle are also important factors to consider, as each radio access technology has a maximum mobility limit. For instance, a vehicle moving at high speed on a freeway would not benefit from scanning a short-range network with a small coverage. It is also not useful to consider networks in the opposite direction of vehicle movement.

The increasing use of small cells has made it difficult to select the most suitable network to achieve optimal handover performance in a heterogeneous multi-tier network environment. This remains an unresolved research problem. In V2I communication, the key characteristic of vehicular networks is high mobility. Therefore, an effective network selection process should consider the vehicle's speed, the direction of movement, and time the vehicle spent within the service area of the selected networks [14][15].

In terms of resource management, many UE served by macro-cell BS will not switch to small cells unless the reference signal of the serving cell falls below a configured threshold. As a result, the resource of small cells may be underutilized while macro-cell BS becomes overloaded. To resolve the issue, some UE in the coverage area in the small cell may be directed to switch to the small cell BS. This can be accomplished by adding a positive bias value to the received power signal from the small cells [16]. D. Jiang *et al.*, [14], proposed a method for selecting a network based on the joint multi-criteria utility under the service area of the selected network, while M. M. Chen *et al.*, [15] said that the vector of motion was used to suggest potential candidate networks based on direction of movement. However, the weakness of both methods is that the resource availability is not checked before network selection, which could increase the probability of a handover drop. To overcome this weakness, Z. Becvar and P. Mach [17] proposed a network selection mechanism where the UE initiates a handover based on the estimated gain in throughput that could be achieved by switching to the target network.

On the other hand, one approach proposed by A. Nadembega *et al.*, [18] uses the historical movement pattern, current path, and digital map to predict the next movement direction and select the best candidate network from those in that direction. This approach requires the UE to store a large database of previously visited BS/AP, which can be a drawback as it requires significant storage capacity and increases equipment costs.

In the coming years, 5G networks will be commercially deployed, consisting of heterogeneous multi-tier and multi-RAT networks that will support a range of diversified services based on V2I communication. To maintain effective QoS and QoE for these services, vehicles will need to quickly and efficiently discover available underlying networks as they move from one location to another, crossing several heterogeneous small networks [19]. However, due to high-speed movement, vehicles may quickly move in and out of the coverage areas of the underlying networks, requiring seamless and fast handover activities. Efficient network discovery is thus a key challenge in V2I communications over 5G networks, as more discovery time leads to

more overall handover delay, deteriorated communication performance, packet loss and potential session disruptions.

In addition to the various standard solutions that allow UE to discover neighboring networks effectively, such as those proposed in [19][20], several researchers have put in significant effort to develop different approaches that enhance the network discovery process energy and time efficiency for the UE. These approaches are aimed at reducing the amount of energy and time the UE spends during the network discovery phase and have been proposed in studies such as in [21] [22]. A recent study proposed an approach to minimize UE power consumption during network scanning for future handover. The author, M. Vondra and Z. Becvar [21] used graph theory to limit the potential candidate networks to only those that share cell boundaries with the serving cells, thus reducing the number of BS/AP that need to be scanned and saving energy. This approach was also proposed by researcher, A. Prasad *et al.*, [22], that consider the use of UE mobility state estimation for small cell discovery in multi-layer networks. They achieved this by prolonging the measurement periods, resulting in infrequent measurements of small cells.

Numerous researchers have worked on solving the problem of network selection for handover in multi-tier heterogeneous environments. Moreover, recent standardization initiatives have also focused on addressing handover issues. To prevent unnecessary handovers, research from H. Ferng and Y. Huang [23] have proposed a network selection mechanism that suggests high-speed vehicles to associate with macro cell eNodeB (eNB) and low-speed vehicles to use microcells based on their estimated coverage area and movement speed, aiming to prevent unnecessary handovers. The study presented by N. Aljeri and A. Boukerche [24] proposed the efficient mobility management of vehicular networks that improved wireless network's performance. A PSO is a computational science approach that optimized a problem by iteratively attempting to improve a candidate solution concerning a given quality measure [25]. M. Al-Gharabally *et al.*, [26] applied a PSO technique to optimize the weights of attributes used in the distance to idea (DIA) method. According to the PSO algorithm, the most suitable neighboring cell for handover was determined to be the one with the highest amount of traffic exchange with the eNB. In conclusion from all the mentioned papers, it shows what have the previous researcher done, what has been achieved and what has not been done yet. Thus, a framework that considers the pre-selection of eNodeB for the handover process is mandatory for the current LTE-A network. An LTE-A network offers much faster peak speeds, for both downloads and uploads, as well as greater reliability. Additionally, LTE-A network allows for global roaming, smooth handover between networks, and interoperability with existing wireless standards.

3. Methodology

This section explains the development of a cell selection framework and describes how the multi-criteria and handover triggering works in LTE-A heterogeneous wireless networks. The cell selection in LTE-A heterogeneous wireless network is shown in Figure 1. The LTE-A multi-tier heterogeneous environment consists of several macro cells and the deployment of small cells.

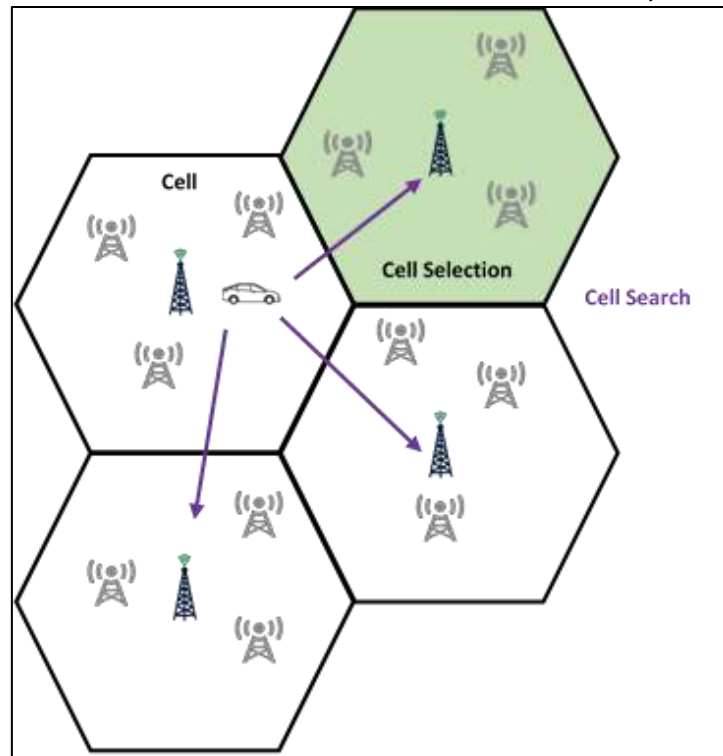


Fig. 1. Cell Selection in LTE-A Heterogeneous Wireless Network

In an LTE-A multi-tier heterogeneous wireless network environment, selecting the best network for handover can be a complex and time-consuming process. When a moving vehicle decides to perform a handover, it must select the best candidate network from a large list of available nearby networks. The proposed solution allows a vehicle with a dual-mode on-board unit (OBU) to self-determine the most appropriate target network for a potential handover based on the calculated network selection probability. This probability is calculated based on two parameters. The first one is the proximity factor, which estimates the geographical distance from the target candidate access point's physical location to the trajectory of the vehicle's movement. The second one is the distance factor, which estimates the remaining distance for the vehicle to reach the coverage area of the target candidate network.

3.1 Proximity Index, ω_m

In this section, a proposed analytical framework for the handover process of UE from the current eNodeB to the target eNodeB is presented in Figure 2. The following mathematical equations [3] have been adopted referencing Figure 2, which illustrates the movement of the UE from the serving eNodeB to another eNodeB, known as the target eNodeB.

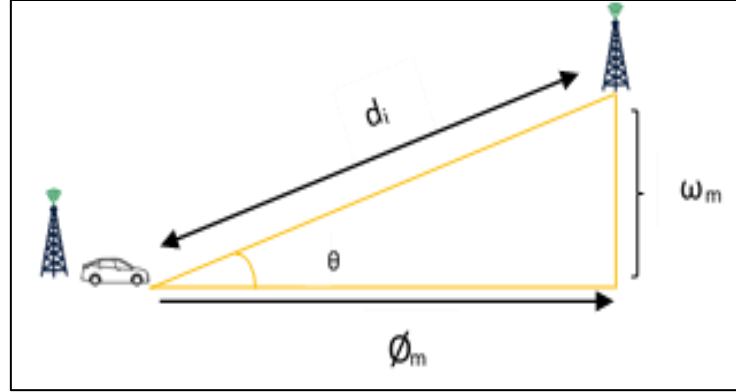


Fig. 2. Analytical eNodeB Selection Framework

The parameters proximity index, ω_m denotes the geo-distance between the physical location of the target candidate eNodeB to the road. The proximity index, ω_m estimated the distance between the geo-location of the target eNodeB based on distance calculation as follows

$$\sin \theta = \frac{\omega_m}{d_i} \quad (1)$$

Rearrange the formula in (1) to get the parameters, ω_m

$$\omega_m = d_i \sin \theta \quad (2)$$

Where d_i is the geo-distance between the current geo-location of UE and the geo-location of target candidate eNodeB. Let (Lat_{UE}, Lon_{UE}) and (Lat_{BS}, Lon_{BS}) denote the coordinate of UE and eNodeB, respectively. ΔLat and ΔLon denote latitude and longitude separation. Given the spherical form of the earth, this geo-distance is calculated by using the Haversine formula [27], which is the function to calculate the great-cycle distance of any two points on the earth. The geo-distance, d_i is calculated as follows

$$haversine(\theta) = haversine(\Delta_{lat}) + \cos(Lat_{UE}) * \cos(Lat_{BS}) * haversine(\Delta_{lon}) \quad (3)$$

Where the angles are in radians and $R = 6371\text{km}$, is the radius of the Earth. Furthermore, the haversine function is given by

$$haversine \delta = \sin^2\left(\frac{\delta}{2}\right) \quad (4)$$

Thus, (3) can be written as:

$$haversine\left(\frac{d_i}{R}\right) = \sin^2\left(\frac{\Delta_{lat}}{2}\right) + \cos(Lat_{UE}) * \cos(Lat_{BS}) * \sin^2\left(\frac{\Delta_{lon}}{2}\right) \quad (5)$$

the distance d_i can be derived from equation (5) as follows

$$d_i = 2R \sin \sqrt{\sin^2\left(\frac{\Delta_{lat}}{2}\right) + \cos(Lat_{UE}) * \cos(Lat_{BS}) * \sin^2\left(\frac{\Delta_{lon}}{2}\right)}$$

(6)

The closer the distance of UE to candidate eNodeB, the greater the value of ω_m . Thus, the better the candidate network and the probability that eNB is added to the candidate list can be formulated as follows

$$P(eNodeB)_{select} = \begin{cases} 1, & \omega_m = 0 \\ \frac{\omega_m}{r}, & 0 < \omega_m < r \\ 0, & \omega_m > r \end{cases}$$

(7)

Where r denotes the coverage area radius.

3.2 Distance Factor, $\emptyset_m$

The distance factor, $\emptyset_m$, gives the estimated remaining distance for the vehicle to reach the coverage area of the target candidate eNodeB. $\emptyset_m$ can be calculated as

$$\emptyset_m = x(eNodeB_{m-UE}) * \cos \alpha$$

(8)

In equation (8), α is the angular movement and it is the angle between the direction of the UE and the UE line of sight with the candidate eNodeB. This implies that the higher the value of $\emptyset_m$, the more time the UE gets to finalize the selection of the candidate eNodeB before it enters the target coverage area. Early selection of the candidate eNodeB for handover can improve the handover performance by reducing the handover latency and minimizing the chances of a failed handover. It also allows for smoother and uninterrupted service for the mobile user during the handover process.

3.3 Handover Triggering By Using Particle Swarm Optimization (PSO) Technique

After determining the candidate eNodeB, the next challenge is to determine the ideal time to initiate the handover process. The underlying assumption in this technique is that the signaling delay, τ associated with handover triggering is constant. Based on this assumption, this section initializes a population array of random positions and velocities. For each particle, the optimum distance value is evaluated by using the PSO technique. The optimum distance in the neighborhood cells is calculated based on the following equation

$$pf = \frac{\cos^{-1}\left(\frac{d}{vt}\right) - \frac{\pi}{2} - \frac{d}{vt}}{\tan^{-1}\left(\frac{a}{d}\right) - \frac{\pi}{2} - \frac{d}{\sqrt{d^2 + a^2}}}$$

(9)

Where pf is the probability of handover failure. The symbol a is the radius of the cell, v is the speed of the UE and d is the optimum distance for handover triggering. The condition of pf is given by

$$pf = \begin{cases} 1 & \tau > \frac{\sqrt{\frac{a^2}{4} + d^2}}{v} \\ p(t < \tau) & \frac{d}{v} < \tau < \frac{\sqrt{\frac{a^2}{4} + d^2}}{v} \\ 0 & \tau \leq \frac{d}{v} \end{cases} \quad (10)$$

The PSO technique is utilized to determine the optimum distance value of d for handover triggering, concerning a given value of pf . Pf is set to 0.02 which means that only two handover attempts will fail out of 100. Another 98% of attempts to handover to target eNodeB will be successful.

4. Results and Discussion

Figure 3 shows the relationship between the proximity index, ω_m and trajectory distance, d_i of the UE. Trajectory movement shows the geo-distance of the current UE location and the geo-location of the target candidate eNodeB. It can be seen that the proximity index, ω_m increases proportionately when the trajectory distance, d_i increases with a particular value of angle, θ where θ is the direction angle.

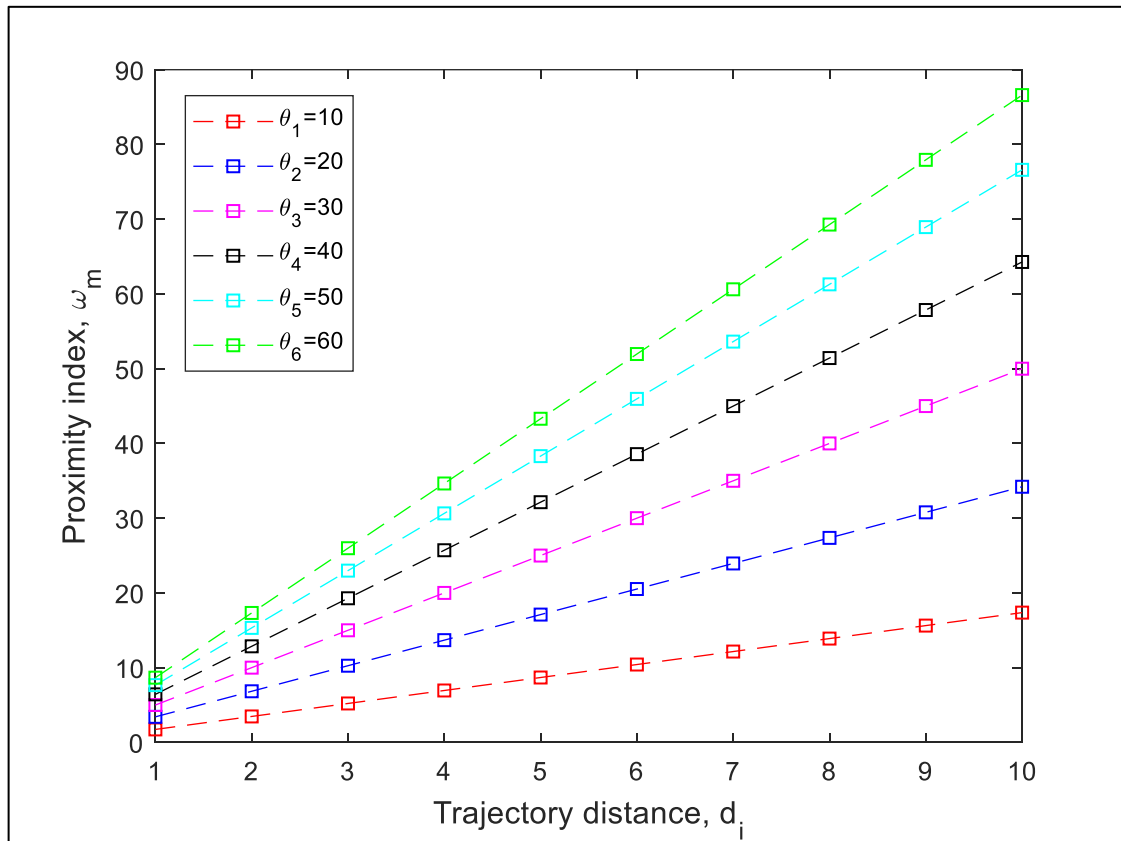


Fig. 3. Relationship Between Proximity Index and Trajectory Distance

The graph shows that the proximity index, ω_m for $\theta = 60^\circ$ is higher than $\theta = 10^\circ$, which means that the larger the angle of UE movement, the greater the value of the proximity index. Higher proximity index shows that a better candidate eNodeB is selected as the shortlisted candidate of the target network.

Figure 4 shows the relationship between the distance factor, Φ_m of UE reaching the coverage area with a distance of eNodeB to the UE, x for a particular value of angle, θ . Figure 4 shows that the distance factor increases as the distance of eNodeB to UE is increased.

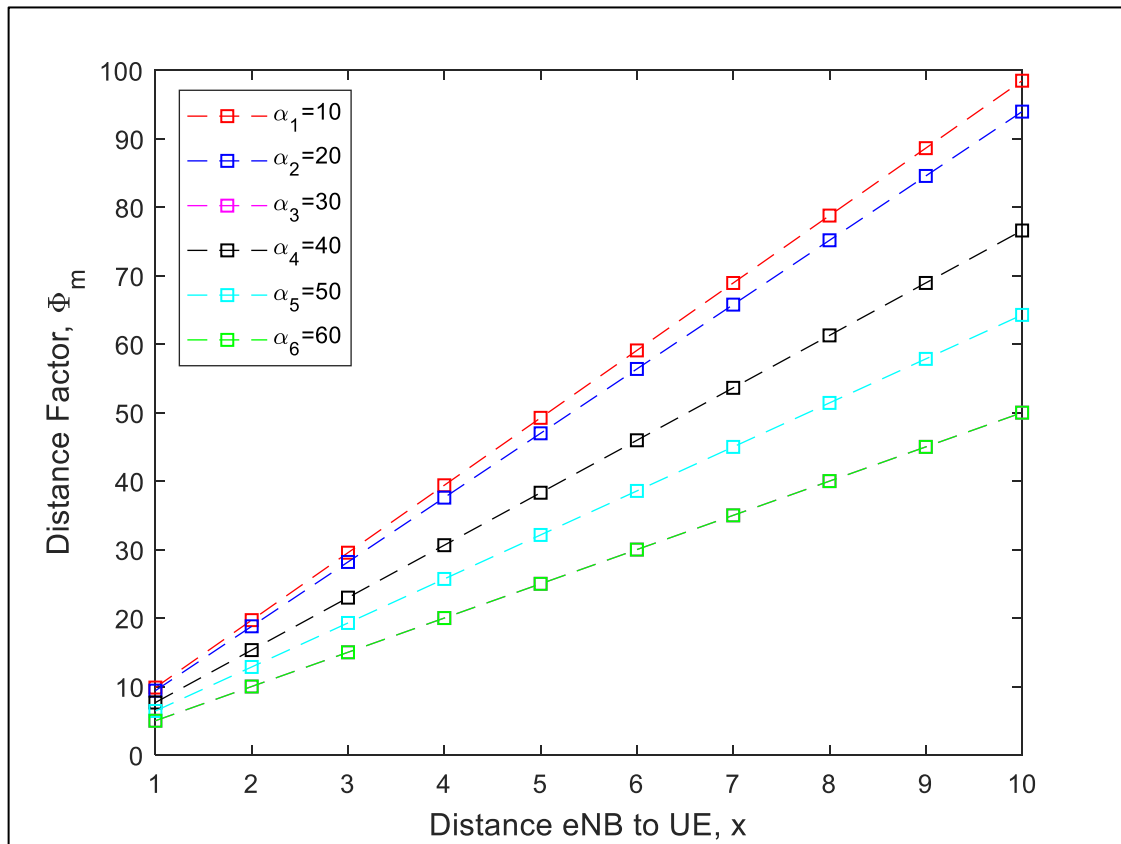


Fig. 4. Relationship Between Distance Factor and The Distance Of eNodeB to UE

From the results, the remaining distance, which is the distance factor, Φ_m is inversely proportional to the angle, α . The lower the value of the angle, α , the higher the value of the distance factor, Φ_m . The higher the value of the distance factor, Φ_m , shows that the more time the UE gets to finalize the selection of the candidate eNodeB. When the angle, α is small, the distance of the eNodeB to the UE is near to be selected as a potential target of eNodeB. Both graphs show that it is appropriate to short-list the candidate eNodeB for the vehicle with reasonable coverage before the handover takes place for an effective communication over multi-tier heterogeneous networks.

The next graph shows the distance for handover triggering, d for high-speed users which is above 40km/h for various time signaling delays, τ as shown in Figure 5. The distance for handover triggering has been optimized using the PSO technique.

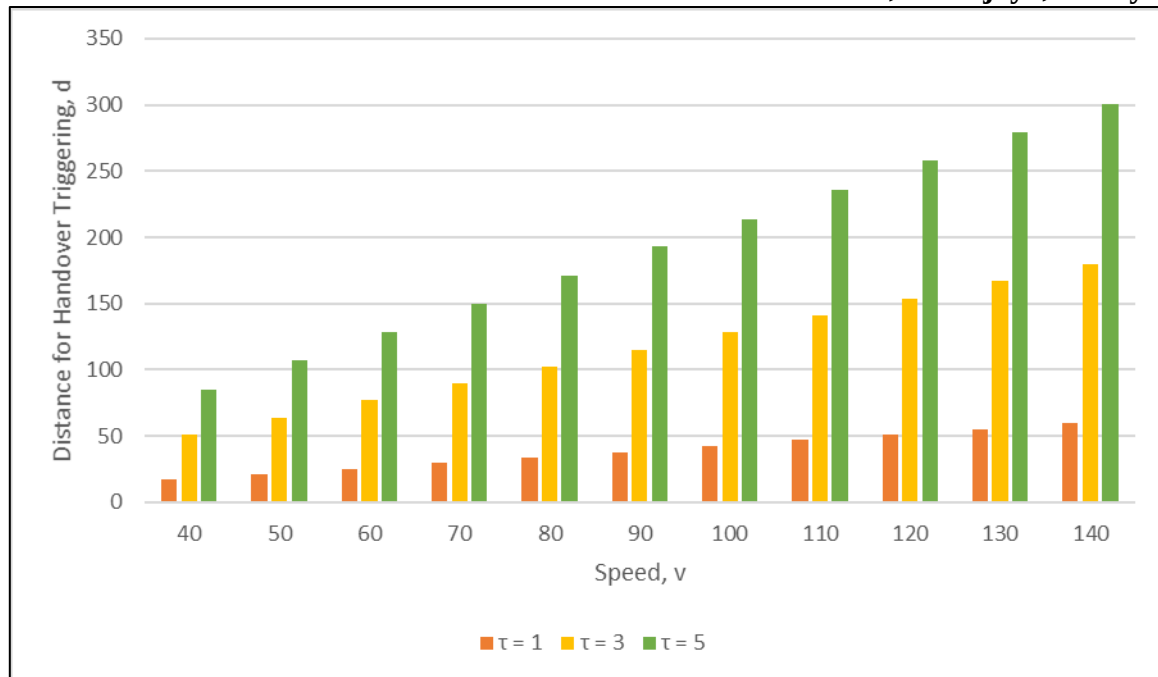


Fig. 5. Graph of Distance for Handover Triggering, D for High-Speed User

From the graph, each bar represents the different distance for the vehicle to initiate handover for three-time signaling delays. As the speed increases, the distance for the UE to initiate handover also increases. It is shown that when the speed of UE is higher, more distance is required for the UE to be handover to the target eNodeB. The graph also shows the higher the time signaling delay, the more distance it will take to handover from the current eNodeB to the target eNodeB. This implies that for the UE moving with high speed, the handover process should be initiated earlier which means more distance is required compared with a slow-moving user.

5. Conclusions

To sum up, this research proposed the parameters for eNodeB selection, which are the movement of UE to target eNodeB with the selected angle of movement that can help in the self-selection of the most appropriate eNodeB for possible handover. The parameters are the proximity factor and distance factor. From the analysis of the results, it shows that the higher the proximity index, the probability to select the eNodeB is higher. The process of selecting the candidate eNodeB for handover at an early stage can lead to the improvements in the overall handover performance by decreasing the handover latency and reducing the probability of a failed handover. Additionally, handover triggering scheme also is proposed by using PSO technique to get the optimum distance to initiate the handover. The result analysis shows that it is essential to choose different distances for handover triggering for different moving UE. Thus, it can be concluded that it is essential to pre-select the target eNodeB that considers the proposed parameters and choose the optimum distance for handover triggering to reduce unsuccessful handovers. For future works, the research will focus on developing an analytical framework to optimize the performance of the handover threshold value.

Acknowledgement

This paper is part of research work that supported by MOHE Grant file no.: FRGS/1/2019/TK04/UITM/02/26 and School of Electrical Engineering, College of Engineering, Universiti Teknologi MARA Shah Alam.

References

- [1] C. Cox, *An Introduction to LTE, LTE-Advanced, SAE, VoLTE and 4G Mobile Communications*, 2nd ed. Wiley & Sons, 2014.
- [2] S. Sesia, I. Toufik, and M. Baker, *LTE - The UMTS Long Term Evolution*, 2nd ed. Wiley & Sons, 2011.
- [3] E. Ndashimye, N. I. Sarkar, and S. K. Ray, "A Network Selection Method for Handover in Vehicle-to-Infrastructure Communications in Multi-tier Networks," *Wirel. Networks*, vol. 26, no. 1, pp. 387–401, 2017.
- [4] N. Algeri and A. Boukerche, "Mobility and Handoff Management in Connected Vehicular Networks," pp. 82–88, 2018.
- [5] S. Kunarak and R. Suleesathira, "Multi-criteria Vertical Handoff Decision Algorithm for Overlaid Heterogeneous Mobile IP Networks," *J. Franklin Inst.*, vol. 357, no. 10, pp. 6321–6351, 2020.
- [6] S. Goudarzi, W. H. Hassan, M. H. Anisi, M. K. Khan, and S. A. Soleymani, "Intelligent Technique for Seamless Vertical Handover in Vehicular Networks," *Mob. Networks Appl.*, vol. 23, no. 6, pp. 1462–1477, 2018.
- [7] A. A. Salih, S. R. M. Zeebaree, M. A. M. Sadeeq, A. S. Abdulraheem, R. R. Zebari, and O. M. Ahmed, "Evolution of Mobile Wireless Communication to 5G Revolution," vol. 62, no. June, 2020.
- [8] K. Samdanis and T. Taleb, "The Road beyond 5G: A Vision and Insight of the Key Technologies," *IEEE Netw.*, vol. 34, no. 2, pp. 135–141, 2020.
- [9] C. Pan, M. El Kashlan, J. Wang, J. Yuan, and L. Hanzo, "User-Centric C-RAN Architecture for Ultra-Dense 5G Networks: Challenges and Methodologies," *IEEE Commun. Mag.*, vol. 56, no. 6, pp. 14–20, 2018.
- [10] C. Pan, H. Zhu, N. J. Gomes, and J. Wang, "Joint Precoding and RRH Selection for User-Centric Green MIMO C-RAN," *IEEE Trans. Wirel. Commun.*, vol. 16, no. 5, pp. 2891–2906, 2017.
- [11] N. Bahra and S. Pierre, "A Hybrid User Mobility Prediction Approach for Handover Management in Mobile Networks," *Telecom*, vol. 2, no. 2, pp. 199–212, 2021.
- [12] S. Sivvam, S. Neeraja, and C. D. Raj, "A Comprehensive Parametric Analysis of Vertical Handoff using Ant Colony Optimization," 2019.
- [13] H. T. Yew, A. Chekima, A. Kiring, A. I. Mbulwa, J. A. Dargham, and S. K. Chung, "RSS Based Vertical Handover Schemes in Heterogeneous Wireless Networks: Past, Present Future," *IEEE Int. Conf. Artif. Intell. Eng. Technol. IICAIET 2020*, no. 1, 2020.
- [14] D. Jiang, L. Huo, Z. Lv, H. Song, and W. Qin, "A Joint Multi-Criteria Utility-Based Network Selection Approach for Vehicle-to-Infrastructure Networking," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 10, pp. 3305–3319, 2018.
- [15] M. M. Chen, Y. Yang, and Z. D. Zhong, "Location-Based Handover Decision Algorithm in LTE Networks Under High-Speed Mobility Scenario," *IEEE Veh. Technol. Conf.*, vol. 2015-Janua, no. January, pp. 1–5, 2014.
- [16] Q. Kuang, J. Belschner, Z. Bleicher, H. Droste, and J. Speidel, "A Measurement-Based Study of Handover Improvement Through Range Expansion and Interference Coordination," *Wirel. Commun. Mob. Comput.*, no. February 2015, pp. 421–430, 2015.

- [17] Z. Becvar and P. Mach, "Mitigation of Redundant Handovers to Femtocells by Estimation of Throughput Gain," *Mob. Inf. Syst.*, vol. 9, no. 4, pp. 315–330, 2013.
- [18] A. Nadembega, A. Hafid, and T. Taleb, "A Destination and Mobility Path Prediction Scheme for Mobile Networks," *IEEE Trans. Veh. Technol.*, vol. 64, no. 6, pp. 2577–2590, 2015.
- [19] K. C. Dey, A. Rayamajhi, M. Chowdhury, P. Bhavsar, and J. Martin, "Vehicle-to-Vehicle (V2V) and Vehicle-to-Infrastructure (V2I) Communication in a Heterogeneous Wireless Network - Performance Evaluation," *Transp. Res. Part C Emerg. Technol.*, vol. 68, pp. 168–184, 2016.
- [20] Y. M. Kwon, J. S. Kim, J. Gu, and M. Y. Chung, "ANDSF-based Congestion Control Procedure in Heterogeneous Networks," *Int. Conf. Inf. Netw.*, pp. 547–550, 2013.
- [21] M. Vondra and Z. Becvar, "Distance-based Neighborhood Scanning for Handover Purposes in Network with Small Cells," *IEEE Trans. Veh. Technol.*, vol. 65, no. 2, pp. 883–895, 2016.
- [22] A. Prasad, O. Tirkkonen, P. Lundén, O. Yilmaz, L. Dalsgaard, and C. Wijting, "Energy-Efficient Inter-Frequency Small Cell Discovery Techniques for LTE-Advanced Heterogeneous Network Deployments," *IEEE Commun. Mag.*, vol. 51, no. 5, pp. 72–81, 2013.
- [23] H. Ferng and Y. Huang, "Handover Scheme with ENode-B Pre-Selection and Parameter Self-Optimization for LTE-A Heterogeneous Network," pp. 10–13, 2016.
- [24] N. Algeri and A. Boukerche, "Mobility Management in 5G-enabled Vehicular Networks : Models , Protocols , and Classification," vol. 53, no. 5, 2020.
- [25] R. Poli, J. Kennedy, and T. Blackwell, "Particle Swarm Optimization : An Overview," *Swarm Intell*, vol. 1, pp. 33–57, 2007.
- [26] M. Al-Gharabally, A. F. Almutairi, and A. A. Salman, "Particle Swarm Optimization Application for Multiple Attribute Decision Making in Vertical Handover in Heterogenous Wireless Networks," vol. 9, no. March, pp. 176–187, 2021.
- [27] M. K. Dahouda, S. Jin, and I. Joe, "Machine Learning-Based Solutions for Handover Decisions in Non-Terrestrial Networks," pp. 1–19, 2023.

PORE PRESSURE DIFFUSION WAVES TRANSMISSION IN OIL RESERVOIR

Mohammed Bashir Abdullahi¹
Shiferaw Regassa Jufar^{1*}
Jang Hyun Lee¹
Tareq Mohammed Al-shami¹
Minh Duc Le²

¹Department of Petroleum Engineering, Universiti Teknologi Petronas, 32610 Seri Iskandar, Perak Darul Ridzuan, Malaysia

²Faculty of Transportation Mechanical Engineering, The University of Danang-University of Science and Technology, Vietnam.

Abstract: Several studies have revealed that the principal wave-induced fluid flow (WIFF) mesoscopic attenuation is due to the compressional (P) wave propagation within oil reservoirs at seismic frequency ranges. The P-wave stimulates the flow of the fluid at mesoscopic-scale heterogeneity that are generally tens of centimeters greater than the dimension of the pore space but shorter than the wavelength. Seismic wave energy is converted to slow Biot's waves, that diffuse outward from the interfaces to maintain pore pressure continuity. Biot's diffusion waves fail to comply with a square-law approach, because of their slow propagation speed and higher attenuation. To characterize the slow Biot's (diffusion) waves in the oil reservoir, a 1D microscopic formulation was used in this study. The result shows that penetration depth and wavelength have identical behavior in different formation permeabilities, and their maximum values move towards low-frequency ranges. However, wave velocity and attenuation have similar trends and their maximum values move towards higher frequency ranges. The pore pressure diffusion process estimated the Biot's diffusion wave's frequency-dependent attenuation. Diffusion waves at a formation layer interface follow the accumulation-depletion relationship instead of the reflection-refraction model. Biot's diffusion waves at the interface of layers are highly attenuated and unable to propagate far away from the source of excitation. Therefore, slow waves can be used to monitor the CO₂ front in oil reservoirs. It can also be used to detect shale gas transport in different layers and identify propagation of fractures during gas/oil recovery, as well as the safe monitoring of the CO₂ storage geo-sequestration process during CCUS projects.

Keywords: Biot's Slow Waves; Oil Reservoir; Pore Pressure; Seismic Waves; Velocity; Wavelength

1. Introduction

Seismic analysis is the primary contributing aspect that the petroleum industry utilizes for exploratory geophysics [1]. Fluid and rock are the two fundamental components that constitute every petroleum reservoir challenges. They are related to two distinct phenomena, namely geomechanics and fluid flow. The appraisal of an oil reservoir must take fluid flow into account. To analyze practically fractured reservoirs and reservoirs that show stress sensitivity, geomechanics is considered to be necessary. Biot was the earliest scientist to introduce the fluid-solid coupling model. Biot's theory and suggested applications are primarily focused on

rock mechanics than the flow of pore fluids. Applying Biot's theory to oil reservoir studies is complicated, especially for engineers who are not particularly interested in geomechanics or fluid-flow processes in poroelastic media [2]. The mechanical behavior of a poroelastic media is influenced by the presence of a flowing viscous fluid. Simultaneously, the viscous pore fluid within the pore spaces alters the internal mechanical condition of the porous skeleton. The poroelasticity model is based on these two related deformation-diffusion processes [1]. Fluid-to-solid coupling occurs when the poroelastic skeleton deforms as a result of a change in the pore fluid pressure or mass. An alteration in the stress of the poroelastic skeleton, also known as solid-to-fluid coupling, can result in an alteration in the pressure of the pore fluid or mass. The fluid-saturated poroelastic material behaves differently over time because of these two processes. Consider compressing the porous media to increase the fluid pressure within the microscopic spaces and allow more fluid to flow. According to Darcy's equation, the poroelastic stresses will become time-dependent due to the diffusive fluid flux and will then react to the pore fluid pressure field. This process has a time-dependent physical description, and once the effects of inertia are ignored, it is simply regarded as quasi-static. The analysis of elastic wave transmission inside poroelastic media is an important subject of research for the petroleum industry due to its importance in seismic prospection and geomechanics [1].

Biot's theory of poroelasticity is typically the basis for models predicting elastic wave propagation in fluid-saturated poroelastic media [1-8]. However, this theory does not adequately account for the substantial attenuation of seismic waves that have been observed [8-9]. One of the possible reasons it's challenging to detect Biot waves on actual rock formations is the limitations imposed by the models' underlying assumptions [4-5]. The Hamiltonian Principle served as the basis for Biot's formulations describing the dynamic motion of pore fluid-saturated poroelastic rock [1,8]. When applying the Hamiltonian Principle to study fluid flow in poroelastic media, there are several restrictions. These restrictions are a reflection of the fundamental idea of Darcy's law. Darcy fluid velocity (superficial velocity) is the flow of pore fluid across an elementary surface in a bulk volume. In actual formation rocks, the pore fluid flows via an incredibly intricate network of many, irregularly shaped pore space (rough surface) channels with different cross-sectional areas. Individual particles of fluid constantly accelerate and decelerate as a result of the intricate shape of the space within the pore. Therefore, there might be a huge difference in the Lagrangian velocities between the individual particles. The mathematical equations defining the elastic displacement of the rock matrix are also developed at the microscopic dimension, where characteristic length has a dimension of a single grain or pore. A parallel network of capillary tubes or simpler poroelastic material arrangement results in less velocity dispersion. It is generally impossible to immediately notice such a huge discrepancy [8]. It is suggested to employ a theoretical framework of slightly compressible fluid flow within a poroelastic media to bridge the gap between the elastic models and length scales of Darcy's law [8]. An equation for parabolic pressure diffusion is developed using a slightly compressible fluid model. Analysis of transient pressure well test data has been established to confirm the model validity [8,10-11]. Due to its underlying presumptions, this technique is only useful for pore pressure fluctuations in the seismic frequency band.

Mandelis research works created the most recent theory of slow Biot's (pore pressure) diffusion waves [12-13]. The concept of pore pressure diffusion waves originated by combining the traditional diffusion equation with an oscillating force (stress) function. These waves are referred to as diffusion or thermal waves due to the pore fluid diffusion and heat conductance phenomena controlled by parabolic equations. A pseudo-wave Helmholtz formulation was

developed and relies on a Fourier transform. Diffusion waves' fundamental feature is that they follow Fick's law of diffusion, which has an impact on how waves behave at formation layers boundaries or interfaces [14]. To clarify some characteristics of heat waves and scattered laser beams moving through a heterogeneous photon density medium, diffusion waves were originally developed [12-13]. Biot's poroelasticity theory anticipated the existence of a slow Biot's compressional (P) wave of the diffusive kind that is greatly attenuated in the seismic frequency band [1-8,12-13]. Because local (squirt) flow develops at the microscopic dimension, classical poroelasticity models fail short to provide a complete explanation [12,15]. Pore fluid pressure buildup and drawdown are related to elastic wave propagation in fluid-saturated poroelastic media [12-13]. The idea of pore pressure diffusion waves in saturated poroelastic media was initially introduced by Norris [16] to characterize pressure relaxation constraints of wave-induced fluid flow (WIFF). Norris [16] focused on the initial studies conducted by Barenblatt *et al.*, [17], in addition to laying the foundations for a pressure diffusion process that can be used in the seismic frequency range associated with pore pressure oscillations [8,12].

Pressure diffusion waves have recently been applied in the geophysical inversion and description of several processes in the petroleum industry, such as deformation of the ground surface caused by carbon dioxide (CO₂) injection into saline aquifers, fault activations, hydraulic fracturing in shale exploration and development, propagation of fractures within sedimentary deposits, subsurface fluid injection-induced micro-seismicity swarms, viscous fingering triggered by pore pressure fluctuations within reservoir during CO₂ sequestration process [12-13]. Research is still ongoing, and it's therefore not clear how Biot's slow diffusion wave (slow speed and heavily attenuated) can be used to interpret and detect fluid flow front in oil reservoirs during seismic enhanced oil recovery (EOR). Due to the continued need for hydrocarbon-based energy sources and the difficulties (both financial and technological) associated with the finding and development of new oilfields, the importance of EOR approaches to enhance oil production from existing reservoirs has increased [18-21]. It is still difficult to predict how seismic excitation for oil recovery will affect the high attenuation behavior of pore pressure waves in the oil reservoir. The objective of this study is to employ pore pressure diffusion waves to determine the fluid flow front during EOR. In this study, the governing equation for the distribution of pore pressure diffusion waves in a poroelastic material is presented together with the Laplace transform periodic wave formulation. The developed model was used to analyze pore pressure fluctuations in oil reservoirs.

2. Methodology

2.1 Pore Pressure Diffusion Waves Physics in The Oil Reservoir

The pore pressure diffusion waves model for oil reservoirs is created when an oscillating force expression and the classical diffusion formulation are coupled [2]. The diffusion expression of pore pressure transmission within saturated poroelastic media is derived when the driving forces are disregarded since there is no delay in the energy conversion mechanism [8,12-13]. We focus on the simple case of pore pressure diffusion waves traveling in a homogenous, 1D poroelastic medium using Chen *et al.*, [2] formulation.

$$\nabla \cdot \left(\frac{\kappa}{\mu} \nabla p \right) + \frac{\kappa}{\mu} c (\nabla p)^2 = \left[\phi c_t + c_b - (1 + \phi) c_s + (c_b - c_s) \frac{d\sigma_m}{dp} \right] \frac{dp}{dt}, \quad (1)$$

The pore fluid pressure meets a decoupled diffusion equation model when the pore fluid is slightly compressible, and the average normal seismic stress remains constant [2]. The

parabolic formulation governs the pore pressure diffusion waves in a poroelastic medium [2,8,12-13]

$$c_t \phi \frac{\partial P(x, t)}{\partial t} = \nabla \cdot \left(\frac{\kappa}{\mu} \nabla P \right), \quad (2)$$

where $P(Pa)$ is the pore pressure, $\mu(Pa \cdot s)$ is the pore fluid viscosity, $c_t(Pa^{-1})$ is the total compressibility coefficient of both pore fluid and the rock matrix, $\phi(\%)$ is the rock porosity, and $\kappa(m^2)$ is the rock permeability.

In Eq. (2), the compressibility of the rock matrix is expressed as a change in porosity caused by variations in fluid pressure. To determine porosity, an elementary volume that is substantially bigger than the volume of a single pore or grain is required. The elementary surface required by Darcy's concept of velocity must also have linear dimensions that are significantly larger than those of a single pore or grain. This makes it possible for the elastic and hydraulic components of Eq. (2) to be stated in the same dimension size, in contrast to the traditional Biot model [8]. If the pressure only depends on one coordinate, as when pressure variations are caused by a source of stimulation within a fracture or at the upper (top) part of the oil reservoir, Eq. (2) can be simplified [2,8,11].

$$c_t \phi \frac{\partial P(x, t)}{\partial t} = \frac{\kappa}{\mu} \frac{\partial^2 P}{\partial x^2}, \quad (3)$$

We can solve Eq. (3) through the application of a Fourier transform [8,12–13]

$$P(x, t) = P_0 \exp(ikx - i\omega t - \alpha x), \quad (4)$$

Eq. (4) represents a plane harmonic wave with frequency $\omega(Hz)$, k is the real part of wave number, α is the attenuation factor in the x -direction. Considering further computations, we can derive the attenuation factor and diffusion wavenumber by inserting Eq. (4) into Eq. (3) [8,12]

$$k = \alpha = \sqrt{\omega \phi \mu c_t / 2\kappa}, \quad (5)$$

The mathematical expression for the wavelength (λ) and phase velocity (V_{ph}) in terms of frequency (ω) is as follows:

$$V_{ph} = \frac{\omega}{k} = \sqrt{2\kappa\omega / \phi c_t \mu}, \quad \lambda = \frac{2\pi}{k} = 2\pi \sqrt{2\kappa / \phi c_t \mu \omega}, \quad (6)$$

2.2 Analytical Solution of Pore pressure diffusion waves in the oil reservoir

This section considers the transmission of pore pressure diffusion waves in a 1D semi-infinitely homogenous poroelastic media. Under appropriate initial and boundary requirements, the parabolic Eq. (2) controls the pressure diffusion wave behavior.

$$t = 0, x > 0, \quad P = 0 \quad t > 0, x = 0, P(0, t) = P_o \quad t > 0, x = \infty, P(\infty, t) = 0, \quad (7)$$

To provide pressure distribution solutions expressed in terms of the complementary error function, *erfc*, we implement the Laplace transform to Eq. (3) using the fundamental concept of the inverse Laplace transform and integrate by partition in the range $(0, \infty)$.

$$P = P_o \operatorname{erfc} \left(\frac{x}{2\sqrt{Dt}} \right), \quad (8)$$

The link between the complementary error function and the error function is as follows:

$$P = P_o \left[1 - \operatorname{erfc} \left(\frac{x}{2\sqrt{Dt}} \right) \right], \quad (9)$$

$$D = \frac{\kappa}{\phi c_t \mu}, \quad (10)$$

The penetration depth (L_ω), which is inversely related to the square root of the diffusion wave frequency (ω), is expressed as follows:

$$L_\omega = \sqrt{\frac{2D}{\omega}}, \quad (11)$$

Eq. (12) [12] can be used to depict the pressure gradient in the x-direction.

$$\frac{\partial P}{\partial x} = P_o \left[-\frac{1}{\sqrt{\pi}\sqrt{Dt}} e^{-\frac{x^2}{4Dt}} \right], \quad (12)$$

3. Results

3.1 Pore Pressure Diffusion Analysis in Oil Reservoir

We evaluate the effects of wave frequency on the attenuation, penetration depth, velocity, and wavelength trends of pore pressure diffusion waves by modeling four conditions with different formation permeabilities of 50md, 100md, 200md, and 450md. A poroelastic medium under consideration is an oil reservoir. Reservoir and fluid properties are provided in Table 1 for the evaluation of Biot's slow diffusion wave.

Table 1: Properties of Rock and Fluid Used in The Investigation of Pore Pressure Wave

Properties	Symbol	Value	Unit
Porosity	ϕ	23	%
Total compressibility	C_t	5.8×10^{-10}	Pa^{-1}
Oil viscosity	μ_o	5.0×10^{-3}	Pa.s

The velocity of the pore pressure wave trends for different formation permeabilities is depicted in Fig. 1 as a function of propagation frequency. When reservoir permeability increases, the maximum values of the pore pressure wave velocity move towards a higher pore frequency value.

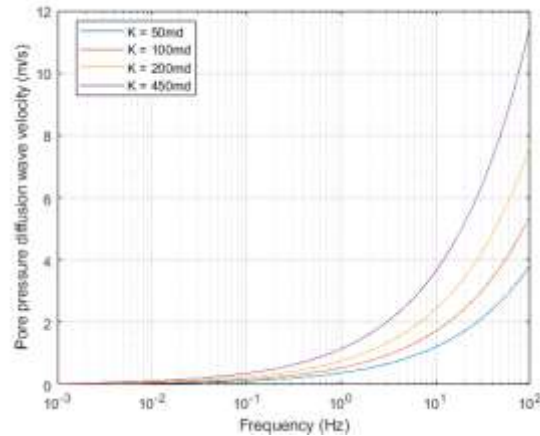


Figure 1: Semilog Graph of Pore Pressure Diffusion Wave Velocity Vs Propagation Frequency for Various Formation Permeabilities.

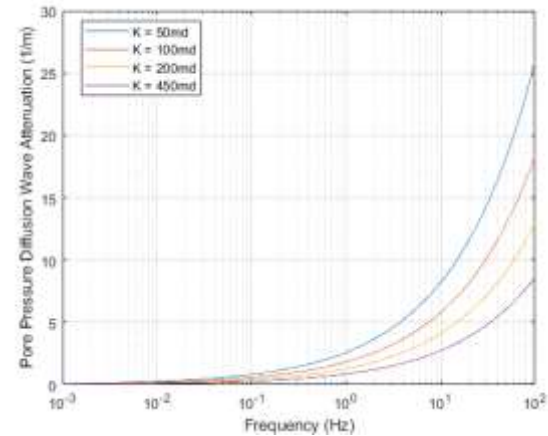


Figure 2: Semilog Graph of Pore Pressure Diffusion Wave Attenuation Vs Propagation Frequency for Various Formation Permeabilities.

Figure 2 demonstrates the relevant frequency-dependent attenuation coefficient. When reservoir permeability rises, the maximum values of the pore pressure wave attenuation component move toward high propagation frequency. So, the attenuation and wave velocity patterns of pore pressure diffusion waves are identical. As a result of considerable energy losses caused by the interaction between pore fluid flow and rock matrix, pore pressure diffusion waves exhibit significant attenuation. Excessive attenuation of seismic waves can be explained by the pore pressure diffusion process, which is not adequately estimated by Biot's poroelasticity law. Pore pressure waves exhibit substantial frequency-dependent attenuation, a slow velocity of dispersion, and a failure to behave in a square-law approach.

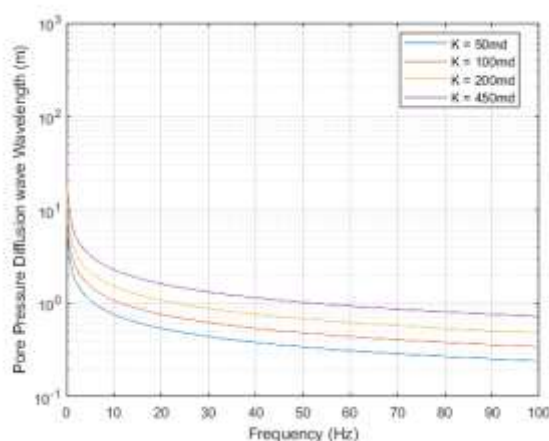


Figure 3: Semilog Graph of Pore Pressure Diffusion Wave Wavelength Vs Propagation Frequency for Various Formation Permeabilities.

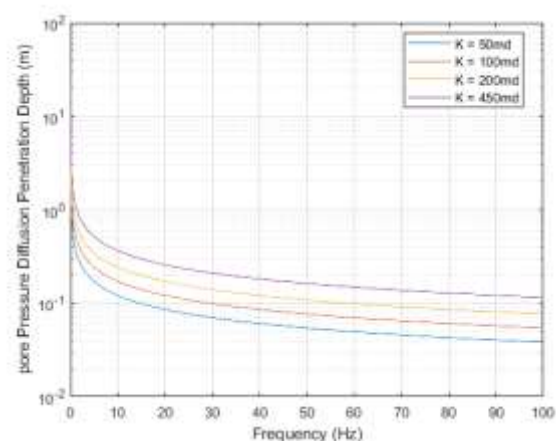


Figure 4: Semilog Graph of Pore Pressure Diffusion Wave Penetration Depth Vs Propagation Frequency for Various Formation Permeabilities.

The wavelength and the depth of penetration are linearly declining functions of wave frequency, as demonstrated in Figs. 3 and 4. Similar behavior in response to permeability variations can be observed in the wavelength and depth of penetration trends generated from the pore pressure diffusion mechanism. When permeability rises, the wavelength and penetration depth maximum values move towards low frequency. At low seismic frequencies, pressure diffusion waves possess high wavelengths and a penetration length. This characteristic facilitates the detection of thin permeable fluid-saturated layers within oil reservoirs.

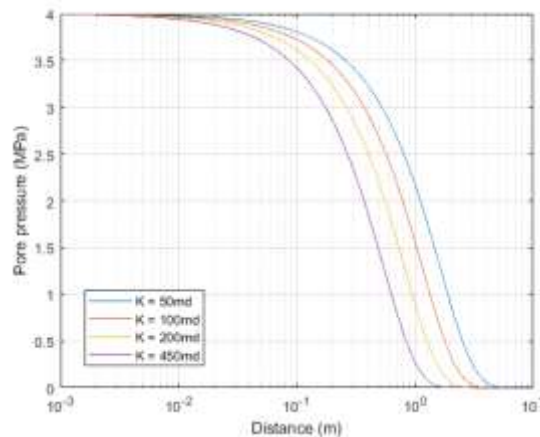


Figure 5: Pore Pressure Propagation Versus Distance for Different Permeabilities At A Travel Time Of 10 S.

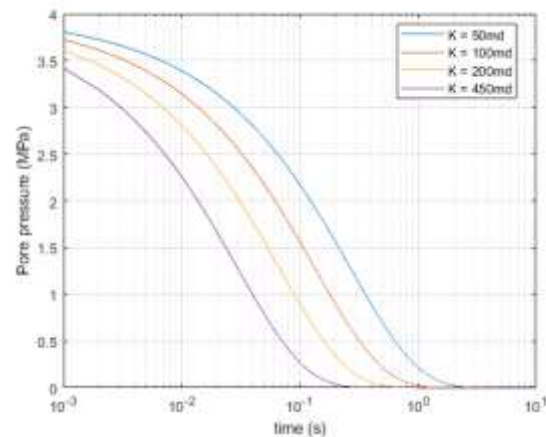


Figure 6: Pore Pressure as A Function Of Travel Time For A Propagated Distance Of 10 M.

The graphs of pore pressure as a function of wave propagation distance for different reservoir permeabilities are presented in Fig. 5. When reservoir permeability declines, the maximum values of pore pressure shift towards longer distances. As the diffusion rate rises, pore pressure also rises for any propagation length. Due to their short propagation distance and high attenuation characteristics at an interface or boundary, pore pressure waves have significant commercial benefits for applications in the detection of CO₂ fronts in depleted oil/gas reservoirs during injection for EOR and monitoring for CCUS projects. Moreover, it can track the migration of shale oil/gas in various formations and use CO₂ flooding to detect the propagation of fractures during oil/gas oil recovery. The pore pressure profile for a 10 m wave propagation distance is shown in Fig. 6 as a function of wave travel time. For a particular formation permeability, pore pressure gradually drops with increasing propagation time. The phase delay in the pore pressure is apparent in Fig. 6.

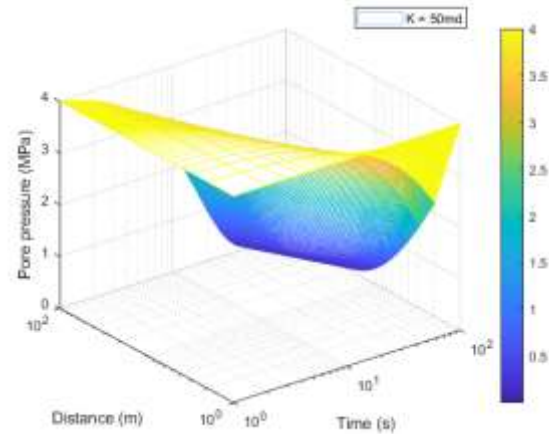
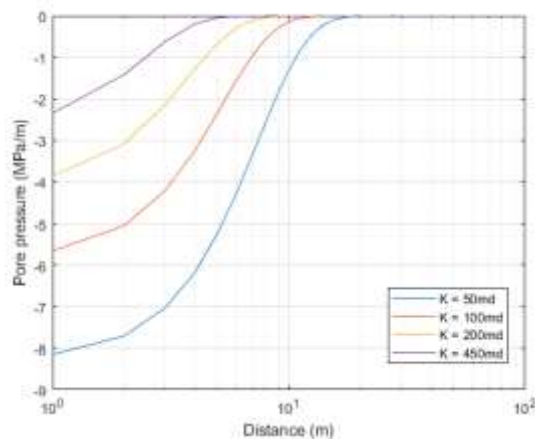


Figure 7: Impacts of Permeability on Figure 8: Presents Pore Pressure Pressure Gradients as A Function of Dispersion in Space And Time Propagation Distances For A 1 S.

Pressure gradients are depicted in Fig. 7 and are represented by a sudden change due to changing formation permeability. As observed in Fig. 8, a recognizable pattern appears for all pressure trends: an initial slow decrease, subsequently followed by an abrupt decline. Despite having limited amplitude at locations far away from the stimulation source, the diffusion waves exhibit a nearly limitless velocity for field transmission. This behavior results in rapid fluctuations across entire pressure fields. A linear law, as opposed to a square law, governs the pressure distribution, which appears as spatial diffusion gradients. For diffusion waves, the penetration length governs the spatially related phase delays. In an isotropic porous medium, pressure diffusion waves lack spatial orientation. The pressure diffusion wave pattern travels in a linear manner, which has an impact on how waves behave at formation layers interfaces. Diffusion waves do not follow the reflection-refraction theory when they come into contact with an interface; instead, they follow the accumulation-depletion theory. The slow slip of natural earthquakes, hydraulic fracturing treatment, shale oil/gas development, and the monitoring of CO₂ geo-sequestration processes for the safety of CCUS projects are significantly impacted by pressure diffusion waves pressure buildup and drawdown.

4. Conclusions

Pressure diffusion waves exhibit strong low-frequency transmission attenuation behaviors, which is very important operationally for locating thin fluid-saturated layers in oil reservoirs. A linear law governs the dynamics of pressure diffusion, which corresponds to the limitless velocity for field transmission, although at far-field points from the stimulation source with little amplitude. Pressure diffusion waves are determined by the accumulation-depletion rule at a mass interaction or formation layers boundaries. Applications for pressure diffusion wave tomography and inverse techniques include the determination of fracture growth during EOR/EGR by CO₂ flooding, monitoring the progression of shale gas mobility in different formations, and the detection of CO₂ front in oil reservoirs during injection for EOR and monitoring for CCUS projects.

Acknowledgment

This work was supported by the YUTP-FRG Project at the Universiti Teknologi PETRONAS with Grant No: 015LC0-447.

References

- [1] David, B.G. "Biot Model for Wave Propagation in Fluid-Saturated Porous Media." *MSc thesis, Imperial College London*, 2015.
- [2] Chen, H.Y. Teufel, L.W. and Lee, R.L. "Coupled Fluid Flow and Geomechanics in Reservoir Study-I. Theory and Governing Equations." *Paper presented at the SPE Annual Technical Conference and Exhibition*, Dallas, Texas, October 1995. doi: <https://doi.org/10.2118/30752-MS>
- [3] Biot, M.A. "General Theory of Three-Dimensional Consolidation." *Journal of Applied Physics* 12, (1941):155- 164.
- [4] Biot, M.A. "Theory of Propagation of Elastic Waves in a Fluid-Saturated Porous Solid. I. Low-Frequency Range." *The Journal of The Acoustical Society of America* 28, no. 2 (1956):168-178.
- [5] Biot, M.A. "Theory of Propagation of Elastic Waves in a Fluid-Saturated Porous Solid. II. Higher Frequency Range." *The Journal of The Acoustical Society of America* 28, no. 2 (1956):179-191.
- [6] Biot, M.A. and Willis, D.G. "The Elastic Coefficients of the Theory of Consolidation." *Journal of Applied Mechanics* 24, (1957):594- 601.
- [7] Biot, M.A. "Mechanics of Deformation and Acoustic Propagation in Porous Media." *Journal of Applied Physics* 33, no. 4 (1962):1482-1498.
- [8] Silin, D. and Korneev, V. "Pressure Diffusion Waves in Porous Media." *SEG Meeting*, 2003.
- [9] Goloshubin, G.M., and Korneev, V.A. "Seismic low frequency effects from fluid-saturated reservoir." *SEG Meeting: Calgary* 2000.
- [10] Barenblatt, G.I. Entov, V.M. and Ryzhik, V.M. 1990. "Theory of Fluid Flows through Natural Rocks." *Dordrecht, Kluwer Academic Publishers*.
- [11] Earlougher, R.C. 1977. "Advances in Well Test Analysis." *New York, Society of Petroleum Engineers*.
- [12] Duoxing, Y. Qi, L. and Lianzhong, Z. "Propagation of pore pressure diffusion waves in saturated porous media." *Journal of Applied Physics* 7, (2015), 117 (13): 134902. <https://doi.org/10.1063/1.4916805>
- [13] Duoxing, Y. Qi, L. and Lianzhong, Z. "Propagation of pore pressure diffusion waves in saturated dual-porosity media (II)." *Journal of Applied Physics* 21, (2016), 119 (15): 154901. <https://doi.org/10.1063/1.4946832>
- [14] Mandelis, A. (2001). "Green Functions and Mathematical Methods of Diffusion-Wave Fields." Springer-Verlag, New York.
- [15] Diallo, M.S. and Appel, E. "Acoustic wave propagation in saturated porous media: reformulation of the Biot-Squirt flow theory." *Journal of Applied Geophysics* 44, (2000): 313–325. [https://doi.org/10.1016/S0926-9851\(00\)00009-4](https://doi.org/10.1016/S0926-9851(00)00009-4).
- [16] Norris, A.N. "Low-frequency dispersion and attenuation in partially saturated rocks." *J Acoust Soc Am* 94, no. (1) (1993): 359–370. <https://doi.org/10.1121/1.407101>
- [17] Barenblatt, G.I., Zheltov, I.P. and Kochina, I.N. "Basic Concepts in the Theory of Seepage of Homogeneous Liquids in Fissured Rocks. " *Journal of Applied Mathematics and Mechanics* 24, (1960):1286-1303. [https://doi.org/10.1016/0021-8928\(60\)90107-6](https://doi.org/10.1016/0021-8928(60)90107-6)
- [18] Abdullahi, M. B., S. R. Jufar, S. Kumar, T. M. Al-shami, and B. M. Negash. "Synergistic effect of Polymer-Augmented low salinity flooding for oil recovery efficiency in Illite-Sand porous media." *Journal of Molecular Liquids* 358 (2022): 119217. <https://doi.org/10.1016/j.molliq.2022.119217>

- [19] Abdullahi, Mohammed Bashir, Kourosh Rajaei, Radzuan Junin, and Ali Esfandyari Bayat. "Appraising the impact of metal-oxide nanoparticles on rheological properties of HPAM in different electrolyte solutions for enhanced oil recovery." *Journal of Petroleum Science and Engineering* 172 (2019): 1057-1068. <https://doi.org/10.1016/j.petrol.2018.09.013>
- [20] Abdullahi, Mohammed Bashir, Shiferaw Regassa Jufar, S. Kumar, Tareq Mohammed Al-shami, and M. D. Le. "Characteristics of SH-wave propagation during oil reservoir excitation using BEM formulation in half-plane model representation." *International Journal of Rock Mechanics and Mining Sciences* 162 (2023): 105303. <https://doi.org/10.1016/j.ijrmms.2022.105303>
- [21] Abdullahi MB, Jufar SR, Dzulkarnain I, Al-shami TM, Le MD (2023b) Seismic Wave Excitation of Mature Oil Reservoirs for Green EOR Technology. *J Adv Res Fluid Mech Thermal Sci* 103(2):180–196. <https://doi.org/10.37934/arfmts.103.2.180196>.

AN EFFECTIVE AIR POLLUTION PREDICTION MODEL USING MACHINE LEARNING ALGORITHMS

Kayalvizhi Subramanian^{1*}
Gunasekar Thangarasu²

¹Faculty of Engineering and Information Technology, MAHSA University, 42610, Jenjarom, Malaysia

²Department of Professional Industry Driven Education, MAHSA University, 42610, Jenjarom, Malaysia

Abstract: Air pollution is a major environmental concern globally, with both developed and developing countries facing its impacts. In recent years, citizens and governments have become increasingly concerned about the effects of air pollution on human health and have proposed sustainable development initiatives to address this issue. In a recent study, air pollution data from the years 2020-2022 was collected from a secondary data source. The data set included six key input features, including SO₂, PM_{2.5}, CO, PM₁₀, NO₂, and O₃ values. To analyse this data, various machine learning models were employed, including linear regression, multiple linear regression, KNN, random forest regression, decision tree regression, support vector regression, and artificial neural networks. To ensure the accuracy of the predictions, mean square error and R square were used to measure the absolute error and forecast precision. Additionally, the importance of each input feature in air pollution was investigated, providing valuable insights into the factors that contribute to air pollution levels. Overall, the use of machine learning algorithms in air pollution estimation and prediction has significant potential to improve our understanding of this critical environmental issue and inform effective strategies for addressing it.

Keywords: Machine Learning, Air Pollution, Prediction, Linear Regression, Artificial Neural Network and KNN

1. Introduction

Air pollution is a serious global issue that has a significant impact on both people and the environment. It occurs when harmful substances are present in the air, leading to negative effects on human health and the natural world. These substances can include various pollutants, such as particulate matter, nitrogen oxides, sulfur dioxide, carbon monoxide, and volatile organic compounds. Air pollution has numerous harmful effects, including respiratory illnesses, heart disease, and stroke, among others. It also harms the environment, contributing to climate change and damaging ecosystems. As such, addressing air pollution is a critical concern for governments, organizations, and individuals alike.

To mitigate the effects of air pollution, various strategies can be employed, such as reducing emissions from vehicles and industrial sources, increasing the use of clean energy sources, and promoting sustainable transportation options. Additionally, monitoring air pollution levels and understanding their sources and impacts is essential for effective policymaking and decision-making. Air pollution is caused by a variety of human activities, such as burning fossil fuels, industrial processes, agriculture and transportation. It is a global problem, affecting not only urban areas but also rural regions. In some cities, air pollution levels have reached dangerous

levels, leading to increased rates of respiratory diseases, heart problems, and even premature death [1].

In addition to its impacts on human health, air pollution can also have negative effects on the natural environment, including acid rain, which can damage crops and forests and contribute to global warming [2]. The global community has taken steps to address this issue, such as signing international treaties and agreements, such as the Paris Agreement, which aims to reduce greenhouse gas emissions and limit the impacts of climate change. However, much work still needs to be done to reduce air pollution and protect the health of individuals and the planet as a whole. This requires a concerted effort from governments, industries and individuals to reduce emission, switch to cleaner energy sources and adopt more sustainable practices.

1.1 Objectives of the Study

Machine learning algorithms can be used to build predictive models that estimate the air pollution level based on various inputs such as weather conditions, emissions data and traffic patterns. These models can provide real-time air quality forecasts, which can help policymakers and individuals make informed decisions.

1.2 Forecasting Air Pollution Using Conventional Methods Can Face Several Challenges

The accuracy of a forecast can be greatly affected by inaccurate or incomplete data. This can be caused by various factors, such as measurement errors, incorrect readings, or missing data, which can result in data quality issues.

Air pollution is a complex phenomenon that is influenced by several factors, including weather patterns, traffic, and industrial activities. These factors often have non-linear relationships with each other, making it challenging for conventional statistical methods to capture and predict accurately. This can result in inaccurate forecasts if these complex relationships are not adequately accounted for.

The levels of air pollution can vary considerably over time and across different locations. To generate accurate forecasts, forecasting models need to consider this temporal and spatial variability. Failure to do so can result in inaccurate predictions that do not reflect the true nature of air pollution.

Conventional statistical methods, such as regression analysis or time series models, have limitations in dealing with complex relationships and large amounts of data. These limitations can result in inaccuracies in the forecast if the model is not configured properly or is not a good fit for the data. Therefore, it is crucial to select appropriate forecasting models that can handle the complexity and volume of data, to ensure accurate predictions.

2. Literature Review

The deterioration of air quality over the past six years has had a negative impact on citizens' quality of life. A recent study [3] investigated the effect of air pollution from vehicles on human health. Improving air quality forecasting is a critical goal for society. Major air pollutants include sulfur dioxide, PM_{2.5}, and nitrogen oxides. Sulfur dioxide, a gas present in the atmosphere, readily combines with other substances to form harmful substances such as sulfuric and sulfurous acids. When inhaled, sulfur dioxide can cause coughing, wheezing, shortness of breath, and tightness in the chest, as well as burning sensations in the nose, throat, and airways. The concentration of sulfur dioxide in the environment can also affect habitat

suitability [4]. This study [5] focused seven regional monitoring areas in geographical and meteorological divergence for forecasting the air pollution results. Besides, they have compared the obtained forecast results with Taiwan, Taipei and London country outputs. Then, investigate the impact of industrial pollution and recommend the improved version of the prediction model to improve their prediction accuracy.

PM2.5, also known as fine particulate matter, is a major health concern when levels in the air are high. A study [6] examined seven regional monitoring areas with diverse geographical and meteorological conditions for air pollution forecasting. The results were compared to outputs from Taiwan, Taipei, and London. The study also analyzed the impact of industrial pollution and proposed improvements to the prediction model to enhance accuracy.

The researchers [7] used the widely popular machine learning technique of support vector regression (SVR) to forecast pollutant and particle levels and the Air Quality Index (AQI). Among the options considered, the radial basis function (RBF) kernel produced the most accurate predictions for SVR. The use of all available variables, rather than just a subset chosen through principal component analysis, was found to be a more effective method. SVR with an RBF kernel accurately predicted hourly pollutant concentrations for substances such as carbon monoxide, sulfur dioxide, nitrogen dioxide, ground-level ozone, and particulate matter 2.5 and PM10.

A different study [8] evaluated the performance of predictive machine learning models for forecasting particulate matter concentration in the air using air quality monitoring data from Taiwan from 2012 to 2017. Mean absolute error (MAE), mean square error (MSE), and coefficient of determination were used as evaluation metrics. Machine learning techniques, such as artificial neural networks (ANNs), are widely used for air quality forecasting due to their ability to take into account a broad range of parameters [9]. Several studies have demonstrated the use of composite models or separate models based on regression algorithms. Other artificial intelligence algorithms include fuzzy logic, generative algorithms, and principal component analysis.

3. Research Methodology

The artificial intelligence techniques have an impact on our daily lives, and artificial intelligence-based algorithms are commonly used for prediction purposes, particularly for air quality forecasting. The process of using machine learning algorithms for air quality index (AQI) prediction typically involves several stages [10]:

3.1 Machine Learning Processing Steps

3.1.1 Data Collection

This involves collecting air quality data such as levels of various pollutants (eg. PM2.5, CO, NO₂, etc.), meteorological parameters (e.g. Temperature, humidity, wind speed, etc.) and other relevant information that can affect air quality.

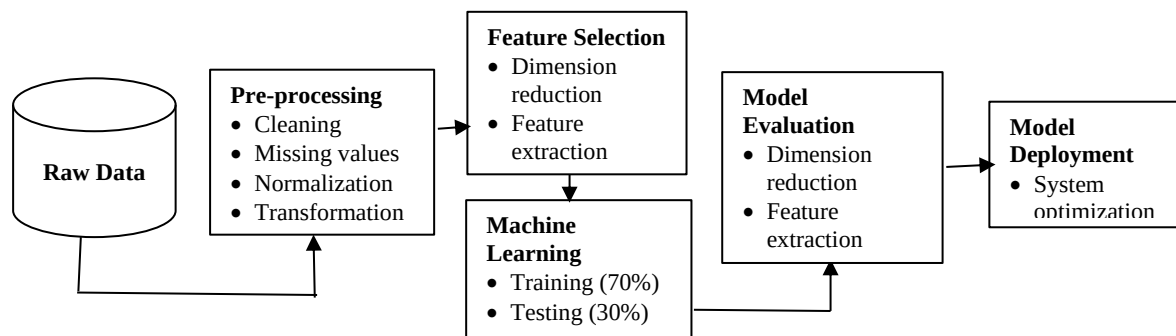


Figure 1: Machine Learning Processing Steps

3.1.2 Data Pre-processing

The collected data may need to be pre-processed to handle missing values, remove outliers, and perform the normalization to ensure the data is ready for analysis.

3.1.3 Feature Engineering

This involves selecting and transforming relevant features of the collected data that can have a significant impact on air quality.

3.1.4 Model Selection

This involves choosing a suitable machine learning model that can be used to predict AQI. Some commonly used models include decision trees, random forests, neural networks, etc.

3.1.5 Model Training

This involves training the selected machine learning model of the pre-processed and engineering data. This stage typically involves optimizing the model's type parameters to obtain the best performance.

3.1.6 Model Evaluation

This involves evaluating the performance of the trained model on a test set, which is a set of data that was not used during the training stage. Common evaluation metrics for AQI prediction include mean absolute error, mean squared error, R-squared, etc.

3.1.7 Model Deployment

Finally, the trained and evaluated model can be deployed in real-world applications, such as predicting AQI in a given location, to help individuals and organizations make informed decisions to mitigate the effects of air pollution. These stages can be iterated multiple times to further improve the model's performance, by fine tuning the features, hyper-parameters or by selecting a different machine learning model

3.2 Data Source

This study collected pollutant concentrations of all elements present in the air to forecast air quality in the National Capital Region (NCR) area covered with 55,083 kilometers around Delhi [11]. The data will be available on the Central Pollution Control Board for the year 2020-2022. We used data from several stations that measure a variety of atmospheric elements. The AQI formulae have been used to calculate the AQI for a specific year using various regression algorithms.

3.3 Air Quality Index Table

An Air Quality Index [12] is used by government agencies to inform the public about how polluted the air is now and how polluted it will become in the future. The risks to public health grow as the AQI rises. Air quality indices differ from country and correspond to different national air quality standards. This study concentrated on the formula used in India to calculate the AQI.

Table 1: Air Quality Index Table in INDIA

AQI Color	Level of Concern	Values of Index
Green	Good	000 -050
Yellow	Moderate	051-100
Orange	Unhealthy for sensitive group	101-150
Red	Unhealthy	151-200
Purple	Very Unhealthy	201-300
Maroon	Hazardous	301 and higher

3.4 Computing AQI

The air quality index is a numerical value used to describe the air quality in a given region. It provides a simple and easy-to-understand measure of air pollution levels, taking into account various pollutants [13].

$$I = \frac{I_{high} - I_{low}}{C_{high} - C_{low}} (C - C_{low}) + I_{low}$$

Where:

C_{low} = The concentration breaking point that is $\leq C$

C_{high} = The concentration breaking point that is $\geq C$

I_{low} = the index breaking point corresponding to C_{low}

I_{high} = the index breaking point corresponding to C_{high}

The AQI is calculated as follows.:

- Determine the concentration of each pollutant in the air. The can be measured by air quality monitoring stations or through remote sensing methods.
- Calculate the pollutant's sub-index for each pollutant by dividing the pollutant concentration by its breakpoint value.
- The AQI is the maximum sub-index value among the pollutants considered.
- Finally, the AQI is categorized into different levels of air quality based on a range o
- AQI values, with higher AQI values indicating poorer air quality.

It is important to note that the AQI calculation method may vary slightly depending on the country or region. However, the basic principle remains the same to provide a simple and easy-to-understand measure of air pollution levels.

3.5 Data Pre-processing

Data preprocessing [14] is a crucial stage in the machine learning process that ensures that the data being used for modeling is in the appropriate format, clean, and consistent. The objective of data preprocessing is to convert raw data into a format that is appropriate for machine learning models.

The common steps involved in data preprocessing in machine learning include:

3.5.1 Data Cleaning

The process of data cleaning includes identifying and correcting any discrepancies or omissions in the data. This may involve eliminating invalid or missing values, or correcting errors in the data set to ensure that it is accurate and complete.

3.5.2 Data Transformation

Data transformation is a crucial step that involves altering the data to make it more appropriate for use in machine learning models. This can include several techniques, such as normalizing the data, converting categorical variables to numerical ones, or scaling the data to a particular range. The aim of this step is to create a more manageable and interpretable data set that can be used to build robust and accurate machine learning models [15].

3.5.3 Data Reduction

Data reduction is the process of reducing the number of variables or features in the data set. This can be done to eliminate irrelevant or redundant information that may hinder model performance, or to simplify the data and make it easier to handle. Data reduction techniques can include feature selection or feature extraction methods, such as principal component analysis (PCA), which can reduce the dimensionality of the data set while retaining important information. The goal of data reduction is to create a more manageable and streamlined data set that is better suited for machine learning modeling.

3.5.4 Data Splitting

Data splitting is a crucial step in machine learning that involves dividing the data set into three distinct subsets: the training set, the validation set, and the testing set. The training set is used to train the model, while the validation set is used to tune the model's hyper parameters and optimize its performance. Finally, the testing set is used to evaluate the model's overall performance and accuracy. The goal of data splitting is to ensure that the model is able to generalize well and perform accurately on unseen data by assessing its performance on a separate data set.

3.6 Training the Model

Data was split 70% for training and 30% for testing. There are numerous models in machine learning for predicting outcomes. Some are in favor of regression, whereas others are in favor of classification. Because we know that AQI has a distinct value, we must train the model using a regression algorithm. The whole processing steps of air quality forecasting shows in Fig.1. The researchers decided to use various algorithms such as support vector regression, multiple linear regression, Lasso-Ridge regression, and k-nearest neighbour, among others.

3.7 Model Evaluation

Evaluating the performance of a machine learning algorithm for forecasting air pollution index (AQI) is an important step to ensure that the model provides accurate and reliable predictions.

4. Result and Discussion

One of the most important goals of the research paper is to develop a machine learning model to predict air pollution. We predict air pollution based on specific data for the year, 2023. We use data from previous years obtained from Central Pollution Control Board website. Through the use of linear regression, Lasso-Ridge regression, KNN, and support vector machine.

Table-2 displays the mean absolute error and root mean squared error for the various machine learning algorithms utilized in the current study. These metrics are valuable for comparing and evaluating different algorithms, and can aid in the selection of the most effective model for a given problem.

Table 2: Root Mean and R Squared Values of Algorithms

Models	Means Absolute Error	Root mean Squared Error
Multiple Linear Regression	39.6776	67.97359
Lasso Regression	39.6587	67.09765
Ridge Regression	39.3942	67.98634
Support Vector Regression	0.09756	0.392473
K-Nearest Neighbor	3.45643	21.86422

Following the evaluation of various types of regression models. The best fit model for predicting AQI is the Support Vector Regression model, which has a 97.5 percent accuracy. This model is the best match. Fig. 2 depicts a comparison of actual and predicted values.

Here, the study shows that almost all values are equal, but some values exhibit anomalous behavior. If we look at the average of AQI month by month in Fig 2, we can easily conclude that AQI is affected by factors other than particle concentration. It concludes that we require more data and add more columns to the dataset of other factors such as temperature and so on.

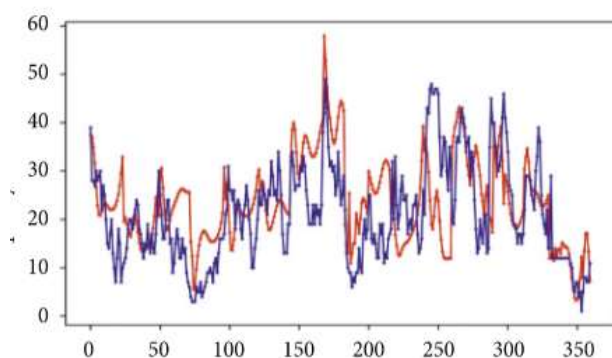


Figure 2: Comparison Between Actual and Predicted Values

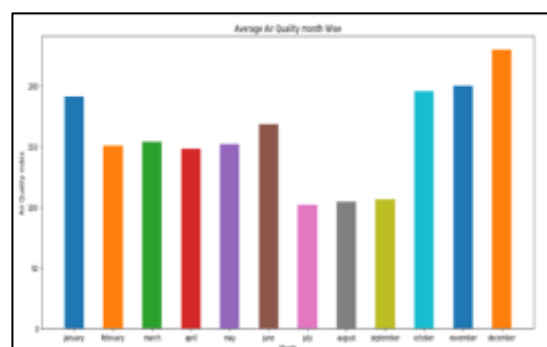


Figure 3: The Average Forecasting Result of AQI

Here can clearly see that the average obtained result for the year 2023 significantly differs from the same months in previous years.

5. Conclusion

It is difficult to predict air quality due to the dynamic nature of pollutants and particulate matter, volatility and wide spatial and temporal variability. At the same time, when it comes to the critical effects of air pollution on people and the environment, the ability to model, predict and monitor air quality is becoming increasingly important, particularly in urban areas. A support vector regression (SVR) study was presented as part of this study to predict pollutant and particulate concentrations and to properly determine the AQI and the method studied produced a suitable model for air pollution, which enabled us to obtain generally effective precision in modelling the concentrations of pollutants such as O₃, CO, and SO₂.

Acknowledgement

This research was not funded by any grant.

Copyright © 2023 Global Academic Excellence (M) Sdn Bhd All Rights Reserved

References

- [1] Srivastava. C, Prakash Singh. A, and Singh. S, "Estimation of Air Pollution in Delhi Using Machine Learning Techniques", *International Conference on Computing, Power and Communication Technologies*, India, 2020, pp.304-309
<https://doi.org/10.1109/GUCON.2018.8675022>
- [2] Aggarwal. P and Jain. S, "Impact of Air Pollutants From Surface Transport Sources on Human Health: A Modelling and Epidemiological Approach", *Environment International*, 2019.vol. 83, pp. 146- 157.
<https://doi.org/10.1016/j.envint.2015.06.010>
- [3] Manuel. I, Mercedes. G and Mahuel. N, "Machine Learning Algorithms to Air Quality Survey", *Artificial Intelligence Review*, 2023, vol. 8, no.09, pp. 367-370.
<https://doi.org/10.1007/s10462-023-10424-4>
- [4] Pandey Gaurav, Bin Zhang, and Li Jian, "Predicting Sub- micron Air Pollution Indicators: A Machine Learning Approach", *Environmental Science: Processes & Impacts*, 2020, pp. 996-1005.
<https://doi.org/10.1039/C3EM30890A>
- [5] Jose Collado and Cristian Pinzon, "Air Pollution Prediction Using Machine Learning Algorithms: A Literature Review," *2022 V Congreso Internacional en Inteligencia Ambiental, Ingeniería de Software y Salud Electrónica y Móvil (AmITIC)*, San Jose, Costa Rica, 2022, pp. 1-6.
<https://doi.org/10.1109/AmITIC55733.2022.9941271>
- [6] Delavar. M, Gholami. A, Reza Shiran. G, Rashidi Gholam Reza, Kurt Fedra, and Afshar. M, "A Novel Method for Improving Air Pollution Prediction Based on Machine Learning Approaches: A Case Study Applied to the Capital City of Tehran", *International journal of Geo-Information*, 2019, vol.8, no. 99.
<https://doi.org/10.3390/ijgi8020099>
- [7] Lee. M, Lin. L, Yuan Chen. C, Yu Tsao, Ting-Hsuan Yao, Min-Han Fei & Shih-Hau Fang, "Forecasting Air Quality in Taiwan by Using Machine Learning", *Scientific reports*, 2020, vol.10, no. 4153, pp. 51-65.
<https://doi.org/10.1038/s41598-020-61151-7>
- [8] Castell. M, Clemente. F. M, Silva. S, and Vanneschi. L, "A Machine Learning Approach to Predict Air Quality in California", *Hindawi, Wiley*, 2020, Vol. 2, pp.1-23.
<https://doi.org/10.1155/2020/8049504>
- [9] Doreswamy, Harishkumar K S, Yogesh K. M, Ibrahim Gad, "Forecasting Air Pollution Particulate Matter (PM_{2.5}) Using Machine Learning Regression Models", *Procedia Computer Science*, 2020, vol.171, pp. 2057-2066.
<https://doi.org/10.1016/j.procs.2020.04.221>
- [10] Jayant Kumar Singh and Amit Kumar Goel, "Prediction of Air Pollution by using Machine Learning Algorithm," *2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS)*, Coimbatore, India, 2021, pp. 1345-1349.
<http://doi.org/10.1109/ICACCS51430.2021.9441902>
- [11] Samaher Al-Janabi, Mustafa. M and Ali Al-Sultan, "A New Method for Prediction of Air Pollution Based on Intelligent Computation", *Soft Computing*, 2020,vol. 24, pp. 661-680,
<https://doi.org/10.1007/s00500-019-04495-1>
- [12] Madhri V. M, Deepak. K, Kamal Agarwal, "Detection and Prediction of Air Pollution using Machine Learning and Deep Learning Techniques", *International Conference on Computing Communication and Intelligent System*, 2023, Greater Noida, India, pp.145-150.
<http://doi.org/10.1109/ICCCIS56430.2022.10037732>

- [13] Zhu. D, Cai. C, Yang. T and Zhou. Z, “A Machine Learning Approach for Air Quality Prediction Model Regularization and Optimization”, *Big Data and Cognitive Computing*, 2018, vol.2, no.5, pp.1-15.
<https://doi.org/10.3390/bdcc2010005>
- [14] Liu. Y, Wang. P, Li. Yong, Wen. L and Deng. X, “Air Quality Prediction Models based on Meteorological Factors and Real-time Data of Industrial Waste Gas”, *Scientific Reports*, 2022, vol. 12, no.22, pp. 2-21.
<https://doi.org/10.1038/s41598-022-13579-2>
- [15] Alexander. B, Yang. Z, “Advances in Air Quality Modeling and Forecasting”, *Global Transitions*, 2020, vol. 2, no.1, pp. 261-270.
<https://doi.org/10.1016/j.glt.2020.11.001>

SOCIAL MEDIA TUNING ATTENTION MODEL FOR INFORMATION SHARING BEHAVIOUR

Zaid Amin¹
Rahma Santhi Zinaida¹

¹Universitas Bina Darma

Abstract: As human beings, we have psychological blind spots as we access the massive abundance of information available to us online from social media and other sources. This onslaught of diverse information can become dissonant in relation to our attention in verifying, critiquing, and believing in the trustworthiness of the information we receive and that is available to us online. An in-depth study is needed in order to describe and model this dissonance regarding our understanding comprehensively. The study in this paper aims to construct and validate a tuning attention model (TAM) that could comprehensively describe users' behavior in sharing information. Our model aims to describe and clarify the integration of our technological solution, which can affect our attention to whether or not we share information. We conducted four studies with 150 participants, including a pilot study where we built ontologies related to theory and constructs. In the second study, we designed the Visual Selective Attention System (VSAS) tool for user experiments. The third study was on behavior change analysis using the Implicit Association Test (IAT). The fourth study involved modeling the overall and the structural model fit and relationship between constructs using Structural Equation Modeling (SEM). This study demonstrated the significance of the detailed processes and relationships of the tuning attention model (TAM) that describes a user's behavior in sharing information online. The result is that the Tuning Attention Model (TAM) shows that all data are significantly valid and reliable in each model relationship. We also prove a significant relationship in the overall and structural model measurement of fit between the constructs, including attention-based design, user attention, social influence, epistemic belief, and user decision-making in sharing information online. These findings can be the basis for designing a range of technological solutions, including the VSAS tool outlined here, which are used to understand and mitigate against the adverse effects of misinformation sharing on social media, particularly prioritizing a humanist approach.

Keywords: Tuning Attention Model (TAM), Information Sharing, User Attention, Misinformation, Social Media

EFFECT OF MATRIX PERMEABILITY ON A HORIZONTAL SHALE GAS WELL PERFORMANCE USING A FULLY COUPLED FLUID FLOW WITH GEOMECHANICS MODEL

Tareq Mohammed Al-Shami¹
Shiferaw Regassa Jufar*¹
Berihun Mamo Negash¹
Mohammed Bashir Abdullahi¹

¹Department of Petroleum Engineering, Universiti Teknologi PETRONAS, Seri Iskandar, Malaysia

Abstract: Recently, hydrocarbon production from unconventional oil and gas resources has emerged. The most common unconventional oil and gas resources are found in shale formations. These shale formations are well-known for being stress-sensitive. Usually, traditional tools are used to model and study production from such formations. However, conventional modeling techniques denote rock deformation using constant rock compressibility. Such an approach is useful for studying conventional hydrocarbon resources where formation stress sensitivity is insignificant. However, when it comes to modeling stress-sensitive formations such as shale formations, it is important to include the rock deformation and its effects on the overall fluid flow in porous media. To consider the rock deformation in the fluid flow model, the coupling of fluid flow with the geomechanics model has to be used. This study utilizes a fully coupled fluid flow with geomechanics model. In addition, shale formations are also known to have an ultra-low matrix permeability, and production usually results from hydraulic fractures that act as flow conduits. Consequently, the effect of matrix permeability on hydrocarbon production is rarely studied. This study focuses on the effect of matrix permeability on the production performance of a single horizontal well in Barnett Shale. It also focuses on the effect of matrix permeability on production performance when the geomechanics effects are coupled with the fluid flow model and decoupled. The results show that the higher the matrix permeability, the better the production performance of the well. The results also show that the higher the matrix permeability, the higher the cumulative production difference between the two models (when the geomechanics effects are coupled and decoupled).

Keywords: Geomechanics Coupling, Shale Gas, Matrix Permeability, Porous Media

1. Introduction

Hydrocarbon remains the primary energy source for most of the world's energy needs [1]–[3]. Hydrocarbon can be gas or oil; its source can be conventional and unconventional [4], [5]. Conventional hydrocarbon formations are usually sandstone or carbonate reservoirs, and the hydrocarbon is usually trapped with a cap rock (e.g., shale formation). Conventional oil and gas reservoirs have relatively high porosity and permeability that can make hydrocarbon extraction achieved using traditional drilling and production methods [6]. The unconventional hydrocarbon resources are usually found in tight sandstone or shale formations. Those formations are well known for having extremely low permeability, which makes hydrocarbon extraction challenging [7]–[9]. However, with the development of technology, such as long horizontal wells and multistage hydraulic fracturing, production from such formations became viable [10]–[13]. As a result, production from unconventional hydrocarbon resources, especially shale formations, increased notably in recent years [14]–[17].

Reservoir simulators are used to numerically study the fluid flow in the reservoir for many purposes, including optimizing reservoir management and production performance. However, the conventional reservoir simulators currently used for such studies utilize constant rock compressibility to denote rock deformation. This assumption can be valid for consolidated rocks that do not deform or have an insignificant deformation due to changes in the reservoir pore pressure, which is the case with conventional reservoirs. However, unconventional reservoirs such as shale are well known for being stress sensitive [18], [19]. This means that the change in the reservoir pore pressure might lead to rock deformation, affecting the fluid flow in the reservoir. For example, the reservoir pore pressure is reduced due to fluid production. The reduction in the reservoir pore pressure increases the effective stress. The increment in the effective stress would lead to rock compaction, which will eventually reduce the permeability. The permeability reduction will change the fluid flow in the reservoir. Consequently, coupling fluid flow with geomechanics in reservoir simulators should be considered for more accurate modeling.

The coupling of geomechanics with fluid flow in porous media can be achieved by applying consolidation theories in reservoir rocks. Terzaghi [20], [21] was the first to introduce the idea of effective stress in consolidation, which provided a framework for fluid and rock interaction. In addition, Biot's theory of poroelasticity is the most widely used theory to explain the poromechanical interaction [22]–[24]. Biot's theory is a mathematical model describing the deformation of a porous medium saturated with fluid. Gretsma [25] was the first to present a unified treatment of rock mechanics in petroleum production engineering. Geomechanics coupling with the fluid flow is important in any numerical modeling affected by poroelasticity, such as CO₂ or H₂ underground storage [26]–[28].

The fluid flow coupling with geomechanics has gained significant attention lately. The effect of the mechanical rock properties and hydraulic fracture geometry on the fluid flow have been studied extensively [18], [29]–[37]. However, when using such a numerical approach, the effect of the reservoir matrix petrophysical parameters is under-investigated. As a result, this paper aims to study the effect of reservoir matrix permeability on the production performance of a gas-producing horizontal well completed with multistage hydraulic fracturing. In addition, this paper aims to show the difference between the cases when the geomechanics effects are coupled with the reservoir simulator and when it is decoupled.

2. Methodology

The present study utilizes a model that fully couples the fluid flow in porous media with geomechanics. The fluid flow model is governed by mass balance (i.e., continuity equation) and assumes a single-phase fluid flow that follows the Darcy law throughout the domain. In addition, the geomechanical deformation is governed by the equilibrium equation. The coupled model assumes the system is isothermal; hence, the temperature will not influence the flow in the model.

2.1 Governing Equations

For a transient flow, the continuity equation for a single-phase gas flow that considers the rock deformation rate is given:

$$\frac{\partial}{\partial t}(\phi \rho_g) + \nabla \cdot (\phi \rho_g (\vec{v}_{gs} + \vec{v}_s)) = \rho_g q_g \quad (1)$$

where, ϕ is the porosity, ρ_g is the gas density, $\vec{v}_g$ is the gas phase interstitial velocity, $\vec{v}_s$ is the deformation rate of the solid phase (rock) due to flow in porous media, and q_g is the sink/source term of the gas phase. In the right-hand side of Eq. (1), the first term is the accumulation, and the second is the fluxes. The left-hand side of Eq. (1) is the sink/source.

Based on Biot & Willis [22] and Geertsma [25] consolidation theory, the coupling between bulk volume and pore volume can be described through the porosity time derivative:

$$\frac{\partial \phi}{\partial t} = \frac{b - \phi}{K_s} \frac{\partial P_t}{\partial t} + (b - \phi) \frac{\partial \varepsilon_v}{\partial t} \quad (2)$$

where b is Biot's coefficient, K_s is the solid grain stiffness (the solid grain bulk modulus). Using the chain rule, the time derivative of the fluid density is expressed by:

$$\frac{\partial \rho_g}{\partial t} = \frac{\partial \rho_g}{\partial P_g} \frac{\partial P_g}{\partial t} = \rho_g c_g \frac{\partial P_g}{\partial t} \quad (3)$$

where c_g is the gas compressibility.

In the continuity Eq. (1), assuming viscous flow and Darcy flow [18], Darcy's law is used in the calculation of the gas flux term [38]. Darcy's law can be described through the gas interstitial velocity as:

$$\vec{v}_g = \phi \vec{v}_{gs} = \frac{\mathbf{k}}{\mu_g} \cdot [-\nabla P_g + \rho_g \vec{g}] \quad (4)$$

where $\mathbf{k}$ is the second-order permeability tensor, μ_g is the gas viscosity, P_g is the pore pressure, and $\vec{g}$ is the gravity vector. In this study the permeability is assumed to be isotropic as a result a single permeability value is used.

The deformation rate of the solid phase (rock) due to flow in porous media is equal to the displacement $\vec{u}$ time derivative:

$$\vec{v}_s = \frac{\partial \vec{u}}{\partial t} \quad (5)$$

where $\vec{u}$ is the solid displacement.

Using small strain approximation, the volumetric strain ε_v is given by:

$$\varepsilon_v = \nabla \cdot \vec{u} \quad (6)$$

Substituting Eqs. (2), (3), (4), (5), and (6) to Eq. (1), and assuming infinitesimal deformation, equation (1) becomes:

$$\rho_g \left(\frac{b-\phi}{K_s} \frac{\partial P_g}{\partial t} + (b-\phi) \frac{\partial \varepsilon_v}{\partial t} \right) + \phi \rho_g c_g \frac{\partial P_g}{\partial t} + \nabla \cdot (\rho_g \vec{v}_g) + \phi \rho_g \frac{\partial \varepsilon_v}{\partial t} = \rho_g q_g \quad (7)$$

Dividing by the gas density term and rearranging:

$$\left(\frac{b-\phi}{K_s} + \phi c_g \right) \frac{\partial P_g}{\partial t} + b \frac{\partial \varepsilon_v}{\partial t} + \nabla \cdot (\vec{v}_g) = q_g \quad (8)$$

Eq. (8) is the governing equation of the model's pressure diffusion (continuity equation) governing the mass balance. Usually, conventional reservoir simulators approximate the velocity field using the multipoint flux approximation (MPFA) method, in which the velocity field can be approximated using the pressure difference between the connected cells. However, the model used in this study treats the velocity as a primary variable, which leads to higher local accuracy of the velocity field [31], [39]. Eq. (4) is the velocity governing equation, which governs the flow in the porous media in the model used in this study.

The equation that governs the rock deformation assuming linear elasticity can be expressed by the Equilibrium equation:

$$\nabla \cdot \sigma + \rho_b \vec{g} = 0 \quad (9)$$

where σ is the stress tensor, ρ_b is the bulk density.

The bulk density is given by:

$$\rho_b = \phi \rho_g + (1-\phi) \rho_m \quad (10)$$

Following the typical geomechanics convention, the compressive stresses are denoted as positive [40], [41], thus the stress tensor is expressed by:

$$\sigma = \sigma_0 - \sigma' + b(P_t - P_0) \mathbf{I} + \rho_b \vec{g} \quad (11)$$

where σ_0 is the initial stress tensor, σ' is the effective stress, P_0 is the initial total pressure and $\mathbf{I}$ is the second rank identity tensor.

For isotropic material the effective stress is given by:

$$\sigma' = 2G\varepsilon + \lambda\varepsilon_v \mathbf{I} \quad (12)$$

where ε is the strain tensor, G is the shear modulus and λ is the first Lamé parameter.

The shear modulus and first Lamé parameter are given in Eqs. (13) and (14), respectively:

$$G = \frac{E}{2(1+\nu)} \quad (13)$$

$$\lambda = \frac{Ev}{(1+\nu)(1-2\nu)} \quad (14)$$

where E is the Young's modulus and ν is the Poisson's ratio.
The strain is related to the solid displacement using the expression:

$$\varepsilon = \frac{1}{2} \left(\nabla \vec{u} + (\nabla \vec{u})^T \right) \quad (15)$$

Eq. (11) is the governing equation for the geomechanical deformation.
Eqs. (4), (8) and (11) are the three governing equations that solve for the pore pressure P_g , gas velocity v_g , and rock displacement $\vec{u}$.

2.2 Numerical Approach

The three governing equations solve for three unknowns: pressure P_h , total velocity $\vec{v}_h$, and solid displacement $\vec{u}_h$. These three variables are interpolated using the mixed finite element scheme. The function spaces used to solve the three governing equations are the Discontinuous Galerkin (DG), the Raviart-Thomas (RT), and the Continuous Galerkin (CG) for the pressure diffusion equation, the total velocity equation, and the solid displacement equation, respectively. The combination of finite element types for these function spaces have been selected because their solution has shown to be stable [31], [42]–[45].

The governing equations' weak form was derived using the traditional Galerkin method. Three test functions were used to derive the weak form equations: P_t , $\vec{V}_t$, and $\vec{U}_t$. Multiplying Eqs. (8), (4), and (11) by the test functions P_t , $\vec{V}_t$, and $\vec{U}_t$, respectively, integrating over the domain Ω and rearranging results into the final weak form of the governing equations:

$$\begin{aligned} \int_{\Omega} \left[\frac{b - \phi^{n+1,m}}{K_s} + \phi^{n+1,m} c_g \right] \frac{P_g^{n+1,m+1} - P_g^n}{\Delta t} P_t d\Omega + \int_{\Omega} b \frac{\varepsilon_v^{n+1,m+1} - \varepsilon_v^n}{\Delta t} P_t d\Omega \\ + \int_{\Omega} \nabla \cdot \vec{v}_t^{n+1,m+1} P_t d\Omega = \int_{\Omega} q_g^{n+1,m} P_t d\Omega, \quad \forall P_t \in M \end{aligned} \quad (22)$$

$$\begin{aligned} \int_{\Omega} \frac{\mu_g}{\mathbf{k}^{n+1,m}} \vec{v}_t^{n+1,m+1} \cdot \vec{V}_t d\Omega = \int_{\Omega} P_g^{n+1,m+1} \nabla \cdot \vec{V}_t d\Omega - \int_{\Gamma} P_g^{n+1,BC} \mathbf{n} \cdot \vec{V}_t d\Gamma \\ + \int_{\Omega} \rho_g \vec{g} \cdot \vec{V}_t d\Omega \quad \forall \vec{V}_t \in Q \end{aligned} \quad (23)$$

$$\begin{aligned} \int_{\Omega} \sigma_0 \cdot \varepsilon(\vec{U}_t) d\Omega - \int_{\Omega} 2G\varepsilon(\vec{u}^{n+1,m+1}) \cdot \varepsilon(\vec{U}_t) d\Omega - \int_{\Omega} (\nabla \cdot \vec{u}^{n+1,m+1}) \cdot \lambda(\nabla \cdot \vec{U}_t) d\Omega \\ + \int_{\Omega} b(P_g^{n+1,m+1}) \cdot (\nabla \cdot \vec{U}_t) d\Omega - \int_{\Omega} b(P_0) \cdot (\nabla \cdot \vec{U}_t) d\Omega \\ + \int_{\Omega} \rho_b^{n+1,m} \vec{g} \cdot \vec{U}_t d\Omega = \int_{\Gamma} (\mathbf{t}^{n+1,BC} \cdot \mathbf{n}) \cdot \vec{U}_t d\Gamma, \quad \forall \vec{U}_t \in U \end{aligned} \quad (24)$$

where the superscripts n and $n+1$ are the previous time step and the current time step, respectively, the superscripts m and $m+1$ are the previous Picard iteration and the current Picard

iteration, respectively, $P_g^{n+1,BC}$ is the stipulated pressure boundary condition, $t^{n+1,BC}$ is the stipulated traction boundary condition.

As a result of gas production, the pore pressure is reduced, thus altering the effective stress, which leads to rock consolidation. Consequently, the porosity and permeability are altered [46]–[50]. When the solid displacement and pore pressure solutions are obtained, the porosity is update using Biot & Willis [22] and Geertsma [25] consolidation theory. Similarly, the permeability is updated using the exponential porosity-dependent permeability developed by Davies and Davies [51].

The initial conditions for the three are unknowns are as below:

$$P_0 = P_i \quad (25)$$

$$v_0 = 0 \quad (26)$$

$$\vec{u}_0 = 0 \quad (27)$$

where P_0 is the pore pressure initial condition, P_i is the initial pore pressure, v_0 is the initial gas velocity, and $\vec{u}_0$ is the initial rock displacement. The flow rate and the rock displacement are assumed to be zero at the beginning of the simulation. The six boundaries shown in figure 1 are no flow boundaries. In addition, the vertical stress S_v is applied at the top boundary as a constant traction t and the other five boundaries are fixed with zero displacement. The sink/source term is defined using the Peaceman equation [52]. The wells are constrained by a constant bottomhole pressure in the sink/source term. It is important to note that the pressure and traction boundary conditions are applied as Neumann boundary conditions, while the total velocity and displacement boundary conditions are applied as Dirichlet boundary conditions.

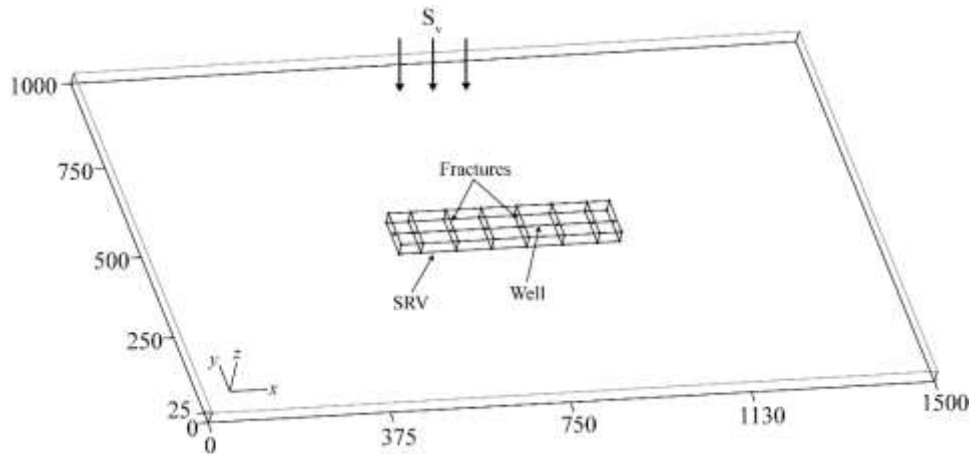


Figure 1: Schematic of the Model Dimensions That Highlights the SRV, the Fractures and the Well

FEniCS Project is used to solve the system of the partial differential equations of the fluid flow coupling with geomechanics [53]. FEniCS Project is an open-source Finite Element library with a high-level Python and C++ interfaces. FEniCS Project uses several advances in automated finite element methodologies, including but not limited to DOLFIN [54], FIAT [55], FFC [56], and UFL [57]. PETSc is the default linear algebra backend for FEniCS Project [58]. For the model used in this study, the weak form equations were coded in Unified Form Language (UFL), while the rest of the model dependencies were coded using Python. The code

used in this study is an extension of the code provided in [45]. FEniCS Project can automatically handle mesh generation, function spaces, and finite element assembly and solution. One of the features that FEniCS Project offers is the use of internal MPI commands that allow for parallel computation [59]. Consequently, the code, including the Picard iterations, can be run in multiple processors, which helps in reducing the computation time significantly.

Equations (22) through (24) are solved monolithically in a mixed finite element scheme. This approach avoids numerical dispersion and unphysical oscillations in the pressure and displacement solutions [43]. Although the monolithic approach (fully coupling) complicates the numerical method by solving long equations, this method can achieve high solution stability and convergence compared to the staggered model (sequential coupling) [60]–[62]. Additionally, FEniCS Project can automatically handle the complications produced by the monolithic approach [45]. The validation of the numerical method used in this study against Terzaghi's one dimensional consolidation theoretical solution has been shown in a previous work of ours [63].

2.3 Data

This study uses data from Barnett Shale to construct the reservoir model. The input used in this study is based on fluid, reservoir, and production data that are available from the literature; Song et al. [64] and Gou et al. [31]. The geomechanics data for the model is based on data from Vermeylen [65], Yu and Sepehrnoori [66], and Zoback [41]. The detailed model data are presented in Table 1. Figure 1 shows the model dimensions of the multistage discrete hydraulic fractures that are used to investigate the performance of the horizontal well. Data in Table 1 is used to match the production history, as shown in figure 2.

Table 1 Fluid, Reservoir, and Rock Parameters for History Matching

Parameter(s)	Value(s)	Unit
Initial reservoir pressure	26.9 (3901.5)	MPa (Psi)
Bottom hole pressure	3.4 (493.13)	MPa (Psi)
Gas Viscosity	2.01×10^{-5} (0.02)	Pa.s (cP)
Matrix permeability	9.87×10^{-19} (1)	m^2 (μD)
Permeability of SRV	4.93×10^{-17} (50)	m^2 (μD)
Matrix porosity	15	%
Porosity of SRV	6.5	%
Fracture conductivity	1.35×10^{-15} (4.5)	$\text{m}^2\text{-m}$ (md-ft)
Fracture half-length	46 (151)	m (ft)
Fracture Spacing	73 (239.5)	m (ft)
Fracture stages	6	-
Bulk density	2.58×10^3 (161.1)	Kg/m^3 (lb/ft ³)
Biot's coefficient	1	-
Young's modulus	40 (5.8×10^6)	GPa
Poisson's ratio	0.25	-
Vertical Stress	44 (6381.6)	MPa (Psi)
Maximum horizontal stress	29 (4206)	MPa (Psi)
Minimum horizontal stress	28 (4061)	MPa (Psi)

3. Results and Discussion

3.1 History Matching

In figure 2, the y-axis represents the gas flow rate at reservoir conditions, while the x-axis represents the production time. Figure 2, shows the history-matching results, the circles represent the data gathered from a gas-producing well in Barnett Shale [64]. The numerical results when the geomechanical effects were decoupled from the reservoir simulators and constant rock compressibility was used instead are represented by the dotted line. At the same time, the numerical results when the geomechanical effects were coupled with the reservoir simulator are represented by the solid line. The production data gathered from the field seem noisy due to unknown production conditions. However, the field data exhibits a clear production profile because a constant bottomhole pressure was used throughout the simulation time. The numerical models (coupled and decoupled) seem to match the production profile of the field production data. While the decoupled numerical model appears to overestimate the production flow rate compared to the field production data, the coupled numerical model matches the field production data more accurately. Since the history matching has been achieved, the numerical analysis is carried out accordingly.

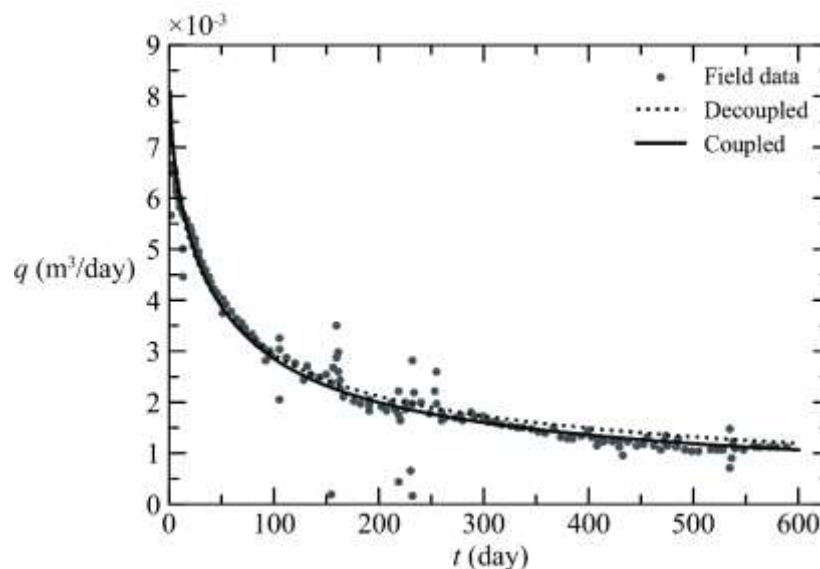


Figure 2: History Matching of the Numerical Model (Coupled and Decoupled) with the Field Data

3.2 Production History

Figure 3(a) shows a comparison of the reservoir production flow rate between the coupled (solid line) and the decoupled (dotted line) cases. In Figure 3(a), the x-axis represents the time, and the y-axis represents the reservoir production flow rate. The production data shown in figure 3(a) showed the reservoir production profile for 600 days. The flow rate resulting from the coupled and the decoupled numerical cases appears to be the same in the first 50 days of production. However, after 50 days of production, there seems to be a slight difference between the coupled and the decoupled cases in the production profile. This difference is due to the consideration of rock deformation in the coupled case. As mentioned in the former sections, the pore pressure is reduced due to production, increasing the effective stress [41], [67]. The increment in the effective stress results in rock compaction, thus reducing the porosity and permeability. This reduction in porosity and permeability is reflected by the slight reduction in

the flow rate when the geomechanical effects were coupled. In Figure 3(a), the difference in flow rate between the coupled and the decoupled cases does not seem to be significant.

Figure 3(b) shows the cumulative production (N_p) for 600 days for the coupled and the decoupled cases. In Figure 3(b), the x-axis represents the time, while the y-axis represents the cumulative production. While the difference in production flow rate was insignificant, as shown in Figure 3(a), the cumulative production that resulted from the coupled and the decoupled cases seem to be more notable, as shown in Figure 3(b). When the geomechanical effects were decoupled, the reservoir simulator overestimated the cumulative production by 11% compared to the case when the geomechanical effects were coupled.

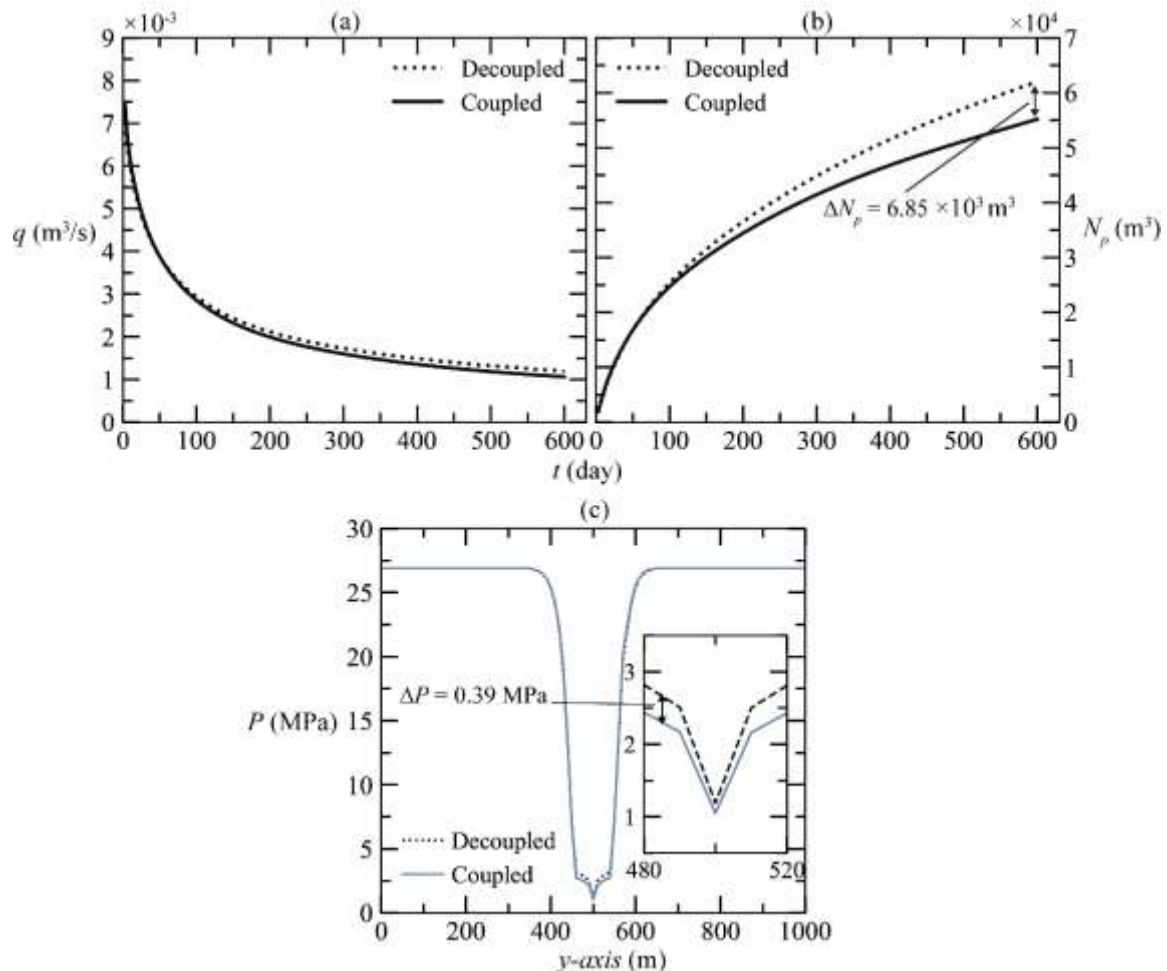


Figure 3: (a) Flowrate for the Coupled vs. Decoupled Cases, (b) a Comparison of the Cumulative Production between the Coupled and the Decoupled Cases, and (c) a Comparison of the Pore Pressure along the Third Fracture between the Coupled and the Decoupled Cases

Figure 3(c) shows the pore pressure along the third fracture along a 1-dimensional line extended in the y direction. The pressure depletion along the third and fourth fractures seems to be higher than that of the other fractures, as shown in figure 4. As a result, the third fracture was chosen for data collection to investigate the pore pressure depletion differences between the coupled and the decoupled cases. In Figure 3(c), the pore pressure for the coupled and the decoupled cases are the same along the third fracture. However, in the middle of the domain, where the well is placed, there appears to be a slight difference. The pore pressure depletion in the coupled case is higher than that of the decoupled case. It is expected that due to the lower

flow rate in the coupled case, the pore pressure would be higher than that of the decoupled case, as shown in Figures 3(a) and 3(b); however, the pore pressure profile shows the opposite of that. This was a direct result of the rock deformation in the coupled case, which reduced the porosity and permeability, resulting in a rapid depletion of the pore pressure.

Figures 4(a) and 4(a*) show a planer top view of the pore pressure contour after 600 days of production for the decoupled and coupled numerical cases, respectively. The pressure contours shown in Figures 4(a) and 4(a*) are cropped from the original model shown in Figure 1, because significant pore pressure depletion occurred mostly at the SRV area, as also shown in Figure 3(c). Both coupled and decoupled cases show the pore pressure depletion propagation in the domain in a ripple effect manner. Notably, the pore pressure reduction along the six fractures seems more prominent than the rest of the SRV area and the reservoir matrix. Generally, the pore pressure contours show a more significant pore pressure depletion when the geomechanical effects are coupled with the reservoir simulator.

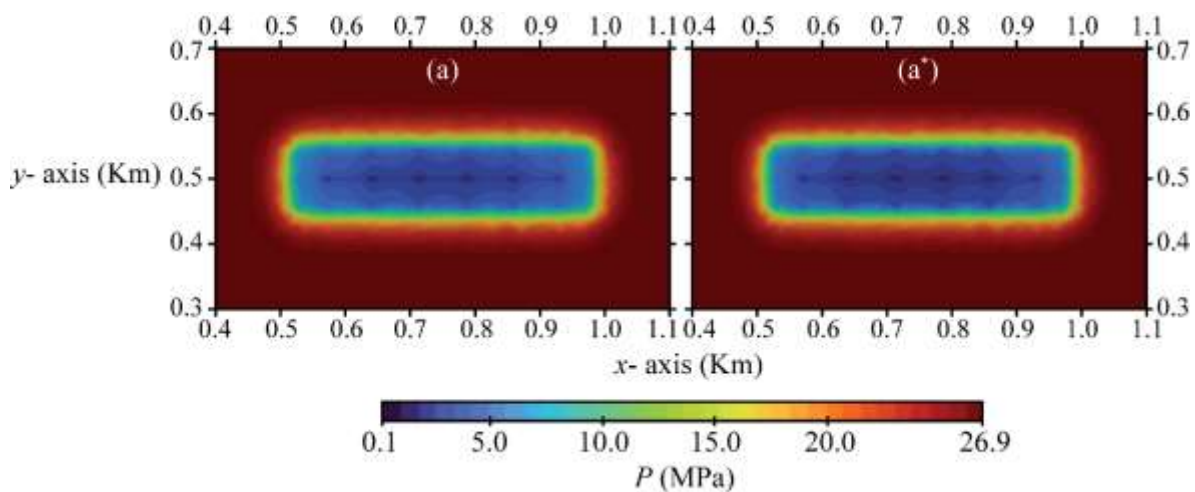


Figure 4: Pore Pressure Contour for (a) the Decoupled Case and (b) the Coupled Case

3.2 Sensitivity Analysis

After the production profile for the decoupled and coupled base cases was shown, this section is aimed to show the sensitivity analysis. The sensitivity analysis carried out in this study investigates the matrix permeability. The varied parameters used for the sensitivity analysis are within the range of the typical shale rock parameters [68], [69]. The reason for that is to study the differences in performance between the decoupled and the coupled cases. Table 2 shows the varied parameters used in this study.

Table 2: The Sensitivity Analysis Parameters

Parameter	Base case	$k \text{ m}^2 (\mu\text{d})$
$k \text{ m}^2 (\mu\text{d})$	9.87×10^{-19}	9.87×10^{-20}
	(1.0)	(0.1)
		1.88×10^{-18}
		(1.9)

Figure 5 shows the sensitivity analysis for the matrix permeability. It shows the cumulative production volume comparison between the coupled and decoupled cases for low and high matrix permeability, as shown in Figures 5(a) and 5(b), respectively. In Figures 5(a) and 5(b),

the x-axis represents the time, while the y-axis represents the cumulative gas production volume. Figure 5 also shows the pore pressure along the third fracture at the end of simulation time (600 days) comparison between the coupled and decoupled cases for low and high matrix permeability, as shown in Figures 5(c) and 5(d), respectively. In Figures 5(c) and 5(d), the x-axis represents the distance along the y-axis, while the y-axis represents the pore pressure at the end of simulation time (600 days).

In Figure 5(a), the solid lines represent the base cases for the coupled and decoupled, similar to Figure 3(b). The dotted lines represent the varied matrix permeability at 0.1 μD . Figure 5(a) shows that as the matrix permeability is reduced, the horizontal well production performance was affected adversely. Figure 5(a) also shows that as the matrix permeability was reduced, the difference between the coupled and the decoupled cases was about half of the base cases. However, the decoupled cases of the low matrix permeability still overestimate the gas cumulative production volume. Figure 5(c) shows the pore pressure along the third fracture, corresponding to the case shown in Figure 5(a). In Figure 5(c), the solid lines represent the base cases for the coupled and decoupled cases, while the dotted lines represent the varied matrix permeability at 0.1 μD . Figures 5(c), 6(a), and 6(a*) show that as the matrix permeability was reduced, the extension of the pore pressure reduction from the SRV was less than that of the base cases. Figure 5(c) also shows that the pore pressure at the center of the fracture (at the well) is lower in the low permeability case compared to the base case in both the coupled and the decoupled cases. In addition, the difference in pore pressure estimation at the center of the third fracture between the coupled and decoupled cases for the low matrix permeability is about half of that of the base cases. Results from Figure 5(a) and 5(c) show that the difference between coupled and decoupled cases is also reduced as the matrix permeability is reduced.

Similar to Figures 5(a) and 3(b), the solid lines in Figure 5(b) represent the gas cumulative production volume for the coupled and decoupled bases cases. In Figure 5(b), the dotted lines represent the coupled and decoupled cases at the matrix permeability of 1.9 μD , which is the high permeability case. The estimated gas cumulative production volume is higher than the base cases for both the coupled and decoupled cases at high matrix permeability, as shown in Figure 5(b). In Figure 5(b), the decoupled case overestimates the gas cumulative production volume, which is the same trend as in Figures 5(a) and 3(b). In addition, the difference between the coupled and the decoupled cases in the high matrix permeability scenario is higher than that of the base cases. Figure 5(d) shows the pore pressure along the third fracture for the case corresponding to the one shown in Figure 5(b). It appears that as the matrix permeability is increased, the pore pressure reduction extends away from the SRV further than that of the base cases, as shown in Figures 5(d), 6(b), and 6(b*). Figure 5(d) also shows that at the center of the third fracture, the pore pressure at the high permeability case was higher than that of the base case at the end of the simulation time (600 days). In addition, the difference between the pore pressure at the center of the third fracture between the coupled and the decoupled cases at high matrix permeability is higher than that of the base cases. This shows that as the matrix permeability gets higher, the decoupled model seems to further overestimate the gas cumulative production volume compared to the lower matrix permeability.

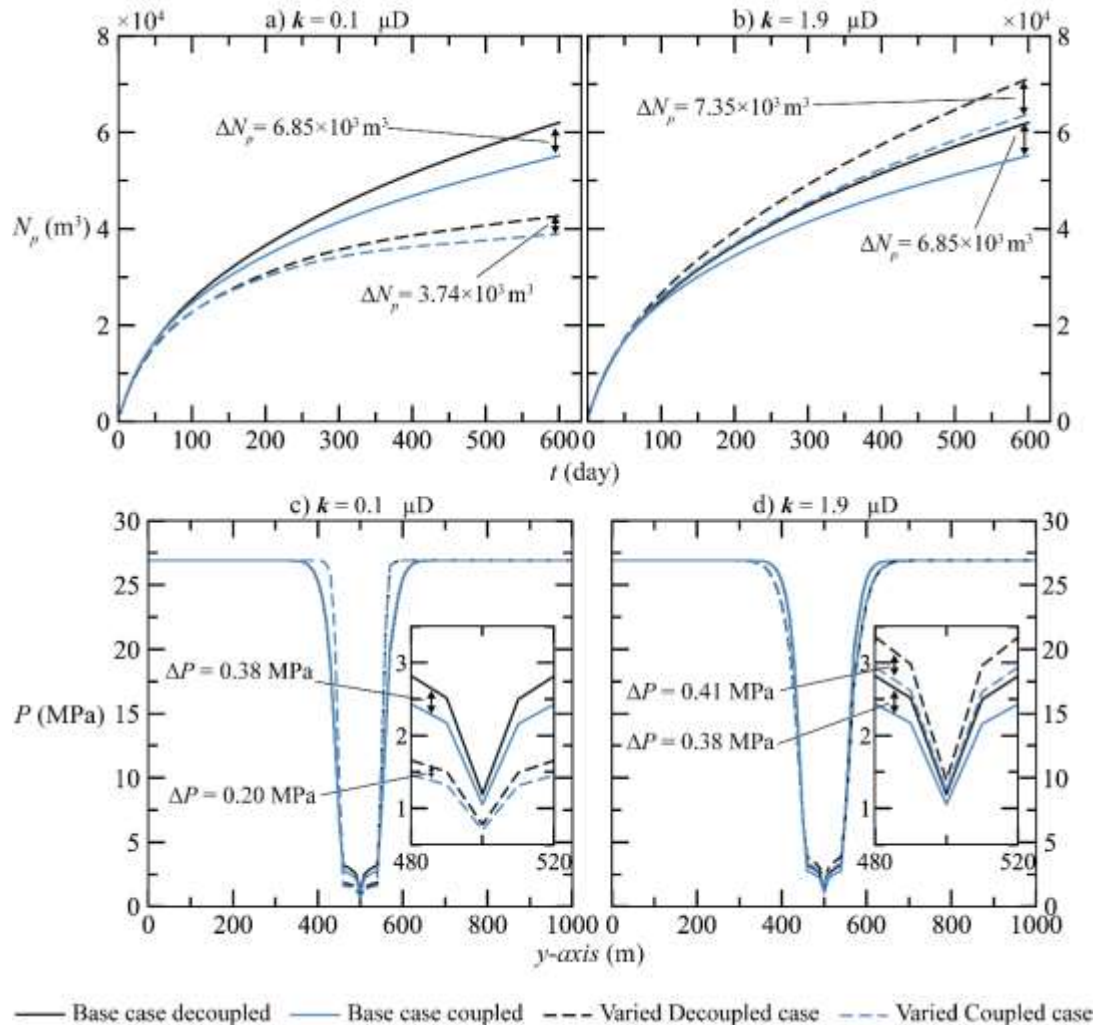


Figure 5: Matrix Permeability Sensitivity Analysis Comparison between Coupled and Decoupled Cases for (a) Cumulative Production for the Low Permeability, (b) Cumulative Production for the High Permeability, (c) the Pore Pressure along the Third Fracture For The Low Permeability Case, (d) The Pore Pressure along the Third Fracture for the High Permeability Case

Figure 6 shows the pore pressure contours of the sensitivity analysis carried out in this study. Figures 6(a) and 6(b) show the pore pressure contours of the decoupled low and high matrix permeability case, respectively. Figures 6(a*) and 6(b*) show the pore pressure contours of the coupled low and high permeability cases, respectively. Figure 6 shows that at low matrix permeability, the pore pressure reduction extension beyond the SRV is limited, as shown in Figures 6(a) and 6(a*). Inversely, at high matrix permeability, the pore pressure contours in figures 6(b) and 6(b*), the pore pressure reduction extends further beyond the SRV. This explains the increased gas cumulative production volume at the high matrix permeability cases. From Figure 6, it can also be noted that the pore pressure reduction inside the SRV is greater at the low matrix permeability cases, as shown in Figures 6(a) and 6(a*). Consequently, the gas is greatly depleted in the SRV when the matrix permeability was low.

In Figure 6, for both the low and high matrix permeability cases, the pore pressure contours show that the pore pressure reduction in coupled cases is greater than that of the decoupled cases. The reason for that is the consideration of the rock deformation in the coupled cases.

According to Biot's theory, the effective stress increases as the pore pressure reduces [22]–[24]. The increment in the effective stress alters the porosity, affecting the permeability. Consequently, the pore pressure contours show that the pore pressure reduction in the coupled cases seems slightly higher than that of decoupled cases. The permeability reduction due to rock deformation led to this greater reduction in the pore pressure in the coupled cases, resulting in a less gas cumulative production volume estimation.

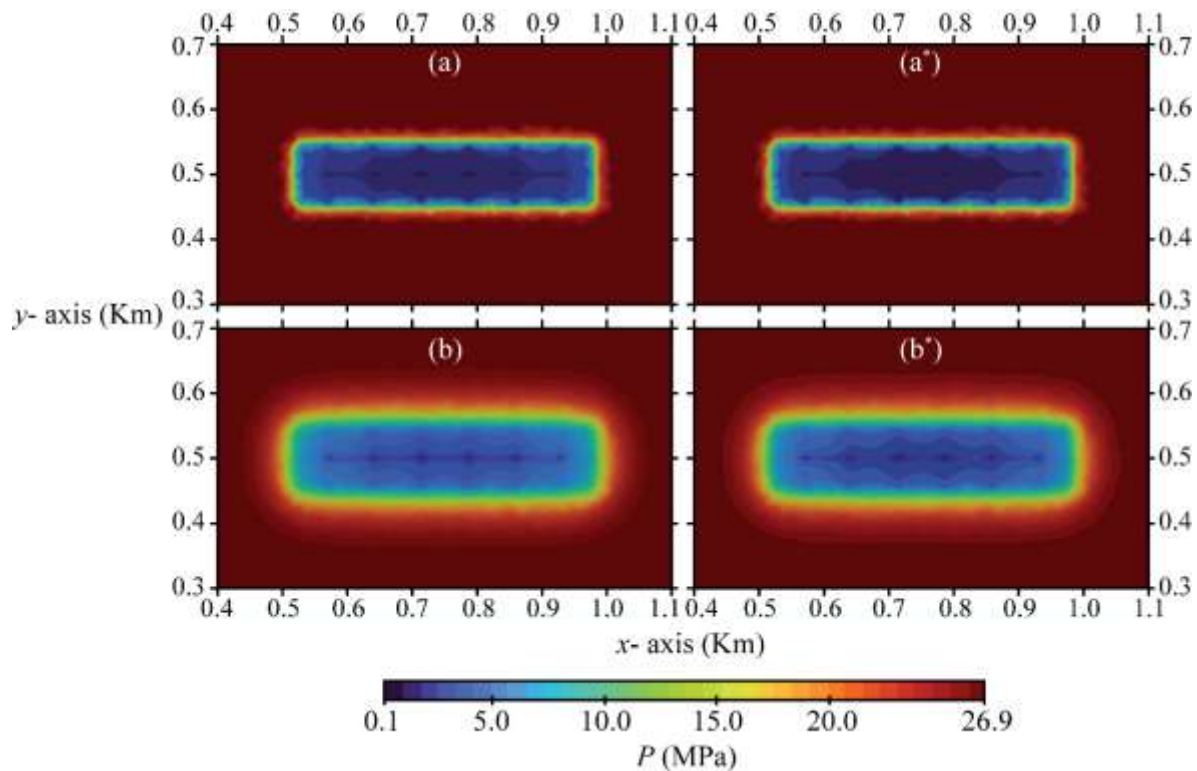


Figure 5: Pressure Contours of Matrix Permeability Sensitivity Analysis for (a) the Low Permeability Decoupled Case and (a*) the Low Permeability Coupled Case, (b) the High Permeability Decoupled Case, and (b*) the High Permeability Coupled Case

4. Conclusions

The utilization of conventional modeling techniques to numerically study hydrocarbon extraction from unconventional resources can lead to inaccurate production estimations. The unconventional hydrocarbon resources are commonly found in shale rock formations. The shale rock formations are well-known for being stress-sensitive. The conventional modeling techniques use a constant value to denote rock deformation. However, for stress-sensitive formation, the mechanical parameters of the rock must be considered for modeling the hydrocarbon extraction from unconventional resources. This paper used a fully coupled fluid flow with geomechanics model to study hydrocarbon production from Barnett Shale. Usually, when such an approach is utilized to study the hydrocarbon production from shale formations, the petrophysical properties, such as the matrix permeability, are rarely studied. This paper studied the effect of matrix permeability on the production performance of a single horizontal well when the geomechanics effect was coupled and decoupled with the fluid flow model. This study showed the difference in production estimation between the cases when the

geomechanics effect was coupled with the fluid flow model and decoupled. The summary of findings from this study is as follows:

- When the matrix permeability was increased, the gas cumulative production volume estimation also increased in the coupled and decoupled cases. In addition, low matrix permeability led to a greater pore pressure reduction in the Stimulated Rock Volume (SRV).
- The decoupled model cases overestimated the gas cumulative production volume for both the high and low matrix permeabilities.
- At low matrix permeability, the difference between the gas cumulative production volume between the coupled and decoupled cases was greatly reduced compared to that of the base case. This means that the decoupled model estimation of the gas cumulative production volume at lower matrix permeability was closer to that of the coupled model.
- At high matrix permeability, the difference between the gas cumulative production volume estimation and the base cases increased. As the matrix permeability increased, the difference between the coupled and decoupled cases in gas cumulative production volume estimation also increased.
- In stress-sensitive formations such as shale formation, although most of the production occurs through hydraulic fractures, the matrix permeability still influences production. The conventional modeling techniques overestimation of the cumulative production increases as the matrix permeability increases.

Acknowledgement

The authors gratefully acknowledge financial support from the YUTP project at the Universiti Teknologi PETRONAS under Grant no. 015LC0-447.

References

- [1] J. I. Mikayilov, S. Mukhtarov, H. Dinçer, S. Yüksel, and R. Aydın, "Elasticity Analysis of Fossil Energy Sources for Sustainable Economies: A Case of Gasoline Consumption in Turkey," *Energies* 2020, Vol. 13, Page 731, vol. 13, no. 3, p. 731, Feb. 2020, doi: 10.3390/EN13030731.
- [2] B. Ovezmyradov and Y. Kepbanov, "Dependence of Central Asian Countries on Fossil Energy and Low Adoption of Non-Hydro Renewables," *SSRN Electronic Journal*, Aug. 2022, doi: 10.2139/SSRN.4193407.
- [3] M. B. Abdullahi, S. R. Jufar, S. Kumar, T. M. Al-shami, and B. . Negash, "Synergistic effect of Polymer-Augmented low salinity flooding for oil recovery efficiency in Illite-Sand porous media," *Journal of Molecular Liquids*, vol. 358, p. 119217, Jul. 2022, doi: 10.1016/J.MOLLIQ.2022.119217.
- [4] Z. Caineng *et al.*, "Conventional and unconventional petroleum 'orderly accumulation': Concept and practical significance," *Petroleum Exploration and Development*, vol. 41, no. 1, pp. 14–30, Feb. 2014, doi: 10.1016/S1876-3804(14)60002-1.
- [5] T. Muther *et al.*, "Unconventional hydrocarbon resources: geological statistics, petrophysical characterization, and field development strategies," *Journal of Petroleum Exploration and Production Technology*, pp. 1–26, Dec. 2021, doi: 10.1007/S13202-021-01404-X/FIGURES/28.
- [6] R. K. Dhir, J. de Brito, R. Mangabhai, and C. Q. Lye, "Use of Copper Slag in Geotechnical Applications," *Sustainable Construction Materials: Copper Slag*, pp. 211–245, Jan. 2017, doi: 10.1016/B978-0-08-100986-4.00006-7.

- [7] H. Abdulelah, B. M. Negash, N. Yekeen, S. Al-Hajri, E. Padmanabhan, and A. Al-Yaseri, "Synergetic Effect of Surfactant Concentration, Salinity, and Pressure on Adsorbed Methane in Shale at Low Pressure: An Experimental and Modeling Study," *ACS Omega*, vol. 5, no. 32, pp. 20107–20121, Aug. 2020, doi: 10.1021/ACSOMEGA.0C01738.
- [8] H. Abdulelah, S. M. Mahmood, and A. Al-Mutarreb, "Effect of Anionic Surfactant on Wettability of Shale and Its Implication on Gas Adsorption/Desorption Behavior," *Energy and Fuels*, vol. 32, no. 2, pp. 1423–1432, Feb. 2018, doi: 10.1021/ACS.ENERGYFUELS.7B03476.
- [9] A. S. Ziarani, R. Aguilera, and A. X. Cui, "Permeability of Tight Sand and Shale Formations: A Dual Mechanism Approach for Micro and Nanodarcy Reservoirs," *Society of Petroleum Engineers - SPE Canada Unconventional Resources Conference 2020, URCC 2020*, Sep. 2020, doi: 10.2118/200010-MS.
- [10] C. JIA, "Breakthrough and significance of unconventional oil and gas to classical petroleum geology theory," *Petroleum Exploration and Development*, vol. 44, no. 1, pp. 1–10, Feb. 2017, doi: 10.1016/S1876-3804(17)30002-2.
- [11] F. JIAO, "Re-recognition of 'unconventional' in unconventional oil and gas," *Petroleum Exploration and Development*, vol. 46, no. 5, pp. 847–855, Oct. 2019, doi: 10.1016/S1876-3804(19)60244-2.
- [12] T. M. Al-Shami, S. R. Jufar, B. M. Negash, and M. B. Abdullahi, "Impact of external excitation on flow behavior of trapped oil blob," *Journal of Petroleum Science and Engineering*, vol. 196, p. 108002, Oct. 2021, doi: 10.1016/j.petrol.2020.108002.
- [13] T. M. Al-Shami, S. R. Jufar, G. M. Hamada, D. A. Kanesan, and U. Djuraev, "Effect of Seismic Excitation on Mobilization of Trapped Oil Globule in Pore Doublet Model," *International Journal of Mechanical Engineering and Robotics Research*, vol. 8, no. 6, pp. 998–1002, 2019, doi: 10.18178/ijmerr.8.6.998-1002.
- [14] Y. Yang and S. Liu, "Review of Shale Gas Sorption and Its Models," *Energy & Fuels*, vol. 34, no. 12, pp. 15502–15524, Dec. 2020, doi: 10.1021/ACS.ENERGYFUELS.0C02906.
- [15] W. Lin, "A Review on Shale Reservoirs as an Unconventional Play – The History, Technology Revolution, Importance to Oil and Gas Industry, and the Development Future," *Acta Geologica Sinica - English Edition*, vol. 90, no. 5, pp. 1887–1902, Oct. 2016, doi: 10.1111/1755-6724.12823.
- [16] H. Abdulelah, B. M. Negash, T. M. Al-Shami, F. A. Abdulkareem, E. Padmanabhan, and A. Al-Yaseri, "Mechanism of CH₄ Sorption onto a Shale Surface in the Presence of Cationic Surfactant," *Energy & Fuels*, vol. 35, no. 9, pp. 7943–7955, May 2021, doi: 10.1021/ACS.ENERGYFUELS.1C00613.
- [17] S. Al-Hajri, B. M. Negash, M. M. Rahman, M. Haroun, and T. M. Al-Shami, "Effect of Silica Nanoparticles on Polymer Adsorption Reduction on Marcellus Shale," *ACS Omega*, vol. 6, no. 44, pp. 29537–29546, Nov. 2021, doi: 10.1021/ACSOMEGA.1C03653/ASSET/IMAGES/LARGE/AO1C03653_0013.JPEG.
- [18] I. Shovkun and D. N. Espinoza, "Coupled fluid flow-geomechanics simulation in stress-sensitive coal and shale reservoirs: Impact of desorption-induced stresses, shear failure, and fines migration," *Fuel*, vol. 195, pp. 260–272, May 2017, doi: 10.1016/j.fuel.2017.01.057.
- [19] C. Wei *et al.*, "A Study of Nonlinear Elasticity Effects on Permeability of Stress Sensitive Shale Rocks Using an Improved Coupled Flow and Geomechanics Model: A Case Study of the Longmaxi Shale in China," *Energies*, vol. 11, no. 2, p. 329, Feb. 2018, doi: 10.3390/en11020329.

- [20] K. Terzaghi, “Die Berechnung der Durchlässigkeitsziffer des Tonen aus Dem Verlauf der Hydrodynamischen Spannungserscheinungen Akademie der Wissenschaften,” *Akademie der Wissenschaften in Wien. Sitzungsberichte*, vol. 11, pp. 105–124, 1923.
- [21] K. Terzaghi, *Erdbaumechanik auf bodenphysikalischer Grundlage (Mechanics of earthworks based on rudiments of soil physics)*. Deuticke, 1925.
- [22] M. A. Biot and D. G. Willis, “The elastic coefficients of the theory of consolidation,” *Journal of Applied Mechanics*, vol. 24, pp. 594–601, 1957.
- [23] M. A. Biot, “Theory of elasticity and consolidation for a porous anisotropic solid,” *Journal of Applied Physics*, vol. 26, no. 2, pp. 182–185, Feb. 1955, doi: 10.1063/1.1721956.
- [24] M. A. Biot, “General theory of three-dimensional consolidation,” *Journal of Applied Physics*, vol. 12, no. 2, pp. 155–164, Feb. 1941, doi: 10.1063/1.1712886.
- [25] J. Geertsma, “Problems of rock mechanics In petroleum production engineering,” in *1st ISRM Congress*, 1966.
- [26] J. Huang and A. Ghassemi, “A poroelastic model for evolution of fractured reservoirs during gas production,” *Journal of Petroleum Science and Engineering*, vol. 135, pp. 626–644, Nov. 2015, doi: 10.1016/J.PETROL.2015.10.007.
- [27] J. Rutqvist, “The geomechanics of CO₂ storage in deep sedimentary formations,” *Geotechnical and Geological Engineering*, vol. 30, no. 3, pp. 525–551, Jun. 2012, doi: 10.1007/S10706-011-9491-0/FIGURES/5.
- [28] A. Elyasi, K. Goshtasbi, and H. Hashemolhosseini, “A coupled geomechanical reservoir simulation analysis of CO₂ - EOR: A case study,” *Geomechanics and Engineering*, vol. 10, no. 4, pp. 423–436, Apr. 2016, doi: 10.12989/GAE.2016.10.4.423.
- [29] L. Y. Chin and L. K. Thomas, “Fully coupled analysis of improved oil recovery by reservoir compaction,” in *Proceedings - SPE Annual Technical Conference and Exhibition*, 1999, vol. 2, doi: 10.2523/56753-ms.
- [30] R. H. Dean and J. H. Schmidt, “Hydraulic-fracture predictions with a fully coupled geomechanical reservoir simulator,” *SPE Journal*, vol. 14, no. 4, pp. 707–714, 2009, doi: 10.2118/116470-PA.
- [31] X. Guo, H. Song, K. Wu, and J. Killough, “Pressure characteristics and performance of multi-stage fractured horizontal well in shale gas reservoirs with coupled flow and geomechanics,” *Journal of Petroleum Science and Engineering*, vol. 163, pp. 1–15, Apr. 2018, doi: 10.1016/j.petrol.2017.12.038.
- [32] J. Jiang and J. Yang, “Coupled fluid flow and geomechanics modeling of stress-sensitive production behavior in fractured shale gas reservoirs,” *International Journal of Rock Mechanics and Mining Sciences*, vol. 101, pp. 1–12, 2018, doi: 10.1016/j.ijrmms.2017.11.003.
- [33] J. Kim, G. J. Moridis, D. Yang, and J. Rutqvist, “Numerical studies on two-way coupled fluid flow and geomechanics in hydrate deposits,” *SPE Journal*, vol. 17, no. 2, pp. 485–501, 2012, doi: 10.2118/141304-PA.
- [34] Y. Liu, L. Liu, J. Y. Leung, and G. J. Moridis, “Sequentially coupled flow and geomechanical simulation with a discrete fracture model for analyzing fracturing fluid recovery and distribution in fractured ultra-low permeability gas reservoirs,” *Journal of Petroleum Science and Engineering*, vol. 189, p. 107042, Jun. 2020, doi: 10.1016/j.petrol.2020.107042.
- [35] S. E. Minkoff, C. M. Stone, S. Bryant, M. Peszynska, and M. F. Wheeler, “Coupled fluid flow and geomechanical deformation modeling,” *Journal of Petroleum Science and Engineering*, vol. 38, no. 1–2, pp. 37–56, May 2003, doi: 10.1016/S0920-4105(03)00021-4.

- [36] M. Moradi, A. Shamloo, and A. D. Dezfuli, "A sequential implicit discrete fracture model for three-dimensional coupled flow-geomechanics problems in naturally fractured porous media," *Journal of Petroleum Science and Engineering*, vol. 150, pp. 312–322, Feb. 2017, doi: 10.1016/j.petrol.2016.12.027.
- [37] W. Yu and K. Sepehrnoori, "Numerical Evaluation of the Impact of Geomechanics on Well Performance in Shale Gas Reservoirs," in *47th U.S. Rock Mechanics/Geomechanics Symposium*, 2013, p. 9, [Online]. Available: <https://doi.org/>.
- [38] M. K. Hubbert, "Darcy's law and the field equations of the flow of underground fluids," *International Association of Scientific Hydrology. Bulletin*, vol. 2, no. 1, pp. 23–59, 1957, doi: 10.1080/02626665709493062.
- [39] D. Yang and M. Sanchez, "A SIMULATOR WITH NUMERICAL UPSCALING FOR THE ANALYSIS OF COUPLED MULTIPHASE FLOW AND GEOMECHANICS IN HETEROGE-NEOUS AND DEFORMABLE POROUS AND FRACTURED MEDIA A Dissertation," 2013.
- [40] D. M. Wood, *Soil behaviour and critical state soil mechanics*. Cambridge University Press, 1990.
- [41] M. D. Zoback, *Reservoir Geomechanics*. Cambridge, New York: Cambridge University Press, 2007.
- [42] C. Zhang, S. J. Zarrouk, and R. Archer, "Development of a fully coupled flow-geomechanics simulator for flow in saturated porous media," in *The International Conference on Computational Methods*, 2015.
- [43] M. Ferronato, N. Castelletto, and G. Gambolati, "A fully coupled 3-D mixed finite element model of Biot consolidation," *Journal of Computational Physics*, vol. 229, no. 12, pp. 4813–4830, Jun. 2010, doi: 10.1016/J.JCP.2010.03.018.
- [44] D. Yang, G. J. Moridis, and T. A. Blasingame, "A fully coupled multiphase flow and geomechanics solver for highly heterogeneous porous media," *Journal of Computational and Applied Mathematics*, vol. 270, pp. 417–432, Nov. 2014, doi: 10.1016/j.cam.2013.12.029.
- [45] R. Haagensohn, H. Rajaram, and J. Allen, "A generalized poroelastic model using FEniCS with insights into the Noordbergum effect," *Computers & Geosciences*, vol. 135, p. 104399, 2020, doi: 10.1016/j.cageo.2019.104399.
- [46] D. J. Soeder and P. L. Randolph, "Porosity, Permeability, and Pore Structure of the Tight Mesaverde Sandstone, Piceance Basin, Colorado," *SPE Formation Evaluation*, vol. 2, no. 02, pp. 129–136, Jun. 1987, doi: 10.2118/13134-PA.
- [47] C. R. McKee, A. C. Bumb, and R. A. Koenig, "Stress-Dependent Permeability and Porosity of Coal and Other Geologic Formations," *SPE Formation Evaluation*, vol. 3, no. 01, pp. 81–91, Mar. 1988, doi: 10.2118/12858-PA.
- [48] J. Geertsma, "The Effect of Fluid Pressure Decline on Volumetric Changes of Porous Rocks," *Transactions of the AIME*, vol. 210, no. 01, pp. 331–340, Dec. 1957, doi: 10.2118/728-G.
- [49] K. K. Wei, N. R. Morrow, and K. R. Brower, "Effect of Fluid, Confining Pressure, and Temperature on Absolute Permeabilities of Low- Permeability Sandstones," *SPE Formation Evaluation*, vol. 1, no. 04, pp. 413–423, Aug. 1986, doi: 10.2118/13093-PA.
- [50] V. M. Dobrynin, "Effect of Overburden Pressure on Some Properties Of Sandstones," *Society of Petroleum Engineers Journal*, vol. 2, no. 04, pp. 360–366, Dec. 1962, doi: 10.2118/461-PA.
- [51] J. P. Davies and D. K. Davies, "Stress-Dependent Permeability: Characterization and Modeling," in *SPE Annual Technical Conference and Exhibition*, 1999, doi: 10.2118/56813-MS.

- [52] A. D. W. Peaceman Member Spe-Aime, "Interpretation of Well-Block Pressures in Numerical Reservoir Simulation(includes associated paper 6988)," *Society of Petroleum Engineers Journal*, vol. 18, no. 03, pp. 183–194, Jun. 1978, doi: 10.2118/6893-PA.
- [53] M. S. Alnaes *et al.*, "The FEniCS Project Version 1.5," *The FEniCS Project Version 1.5*, vol. 3, no. 100, pp. 9–23, Dec. 2015, doi: 10.11588/ans.2015.100.20553.
- [54] A. Logg and G. N. Wells, "DOLFIN: Automated finite element computing," *ACM Transactions on Mathematical Software*, vol. 37, no. 2, pp. 1–28, Apr. 2010, doi: 10.1145/1731022.1731030.
- [55] R. C. Kirby, "Algorithm 839," *ACM Transactions on Mathematical Software (TOMS)*, vol. 30, no. 4, pp. 502–516, Dec. 2004, doi: 10.1145/1039813.1039820.
- [56] R. C. Kirby and A. Logg, "A compiler for variational forms," *ACM Transactions on Mathematical Software (TOMS)*, vol. 32, no. 3, pp. 417–444, Sep. 2006, doi: 10.1145/1163641.1163644.
- [57] M. S. Alnaes, A. Logg, K. B. Oelgaard, M. E. Rognes, and G. N. Wells, "Unified Form Language: A domain-specific language for weak formulations of partial differential equations," *ACM Transactions on Mathematical Software*, vol. 40, no. 2, Nov. 2012, Accessed: 31-Mar-2021. [Online]. Available: <http://arxiv.org/abs/1211.4047>.
- [58] S. Abhyankar, J. Brown, E. M. Constantinescu, D. Ghosh, B. F. Smith, and H. Zhang, "PETSc/TS: A Modern Scalable ODE/DAE Solver Library," Jun. 2018, doi: 10.48550/arxiv.1806.01437.
- [59] E. Gabriel *et al.*, "Open MPI: Goals, concept, and design of a next generation MPI implementation," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 3241, pp. 97–104, 2004, doi: 10.1007/978-3-540-30218-6_19/COVER.
- [60] A. E. Kolesov, P. N. Vabishchevich, and M. V. Vasilyeva, "Splitting schemes for poroelasticity and thermoelasticity problems," *Computers & Mathematics with Applications*, vol. 67, no. 12, pp. 2185–2198, Jul. 2014, doi: 10.1016/J.CAMWA.2014.02.005.
- [61] J. Kim, H. A. Tchelepi, and R. Juanes, "Stability and convergence of sequential methods for coupled flow and geomechanics: Fixed-stress and fixed-strain splits," *Computer Methods in Applied Mechanics and Engineering*, vol. 200, no. 13–16, pp. 1591–1606, Mar. 2011, doi: 10.1016/J.CMA.2010.12.022.
- [62] T. M. Al-Shami *et al.*, "A comprehensive review of interwell interference in shale reservoirs," *Earth-Science Reviews*, vol. 237, p. 104327, Feb. 2023, doi: 10.1016/J.EARSCIREV.2023.104327.
- [63] T. M. Al-Shami, S. R. Jufar, M. B. Abdullahi, and B. M. Negash, "A Study on the Production Performance of a Horizontal Shale Gas Well Using a Fully Coupled Flow and Geomechanics Modeling," *Journal of Advanced Research in Fluid Mechanics and Thermal Sciences*, vol. 105, no. 1, 2023, doi: Accepted.
- [64] H. Song, Z. Li, Y. Wang, W. Zhu, Y. Cao, and J. Killough, "A Novel Approach for Modeling Gas Flow Behaviors in Unconventional Reservoirs with Nanoporous Media," 2015, doi: 10.2523/iptc-18316-ms.
- [65] J. P. Vermynen, "Geomechanical studies of the Barnett shale, Texas, USA," Stanford University, 2011.
- [66] W. Yu and K. Sepehrnoori, "Simulation of gas desorption and geomechanics effects for unconventional gas reservoirs," *Fuel*, vol. 116, pp. 455–464, Jan. 2014, doi: 10.1016/j.fuel.2013.08.032.
- [67] F. O. Ajisafe, I. Solovyeva, A. Morales, E. Ejofodomi, and M. M. Porcu, "Impact of

Well Spacing and Interference on Production Performance in Unconventional Reservoirs, Permian Basin,” in *Unconventional Resources Technology Conference*, 2017, p. 16, doi: 10.15530/URTEC-2017-2690466.

- [68] C. Guo, M. Wei, and H. Liu, “Modeling of Gas Production from Shale Reservoirs Considering Multiple Transport Mechanisms,” *PLOS ONE*, vol. 10, no. 12, p. e0143649, Dec. 2015, doi: 10.1371/journal.pone.0143649.
- [69] Y. Pei and K. Sepehrnoori, “Investigation of Parent-Well Production Induced Stress Interference in Multilayer Unconventional Reservoirs,” *Rock Mechanics and Rock Engineering*, pp. 1–22, Jan. 2022, doi: 10.1007/S00603-021-02719-1.

POLYETHERSULFONE – POLYITACONIC ACID (PES-PIA) MEMBRANE FOR HIGHLY SELECTIVE DYES REMOVAL

Uswatun Hasanah Abdul Roni¹
Mohd Hafez Mohd Isa^{1*}

¹ Faculty of Science and Technology, Universiti Sains Islam Malaysia, Bandar Baru Nilai, 71800 Nilai, Negeri Sembilan, Malaysia

Abstract: The hazardous and mutagenic properties of synthetic dyes effluent discharge into water bodies have led to the development of membrane separation for water treatment. pH-sensitive membranes have been widely employed as they exhibited remarkable improvements in the permeability and separation performance of PES membranes. In this study, the polyethersulfone (PES) membrane incorporated with polyvinylpyrrolidone (PVP), pluronic F108, and polymerized itaconic acid by phase inversion were fabricated and characterized. Polymerized itaconic acid (PIA) has been employed in the PES dope solution via blending technique. Fabricated membrane with different weight percentages of PIA were studied. The dye removal capacity, chemical composition, and thermal stability were characterized using UV-Vis spectrometers, fourier transform infrared (FTIR), and thermogravimetric analyzer (TGA) respectively. The study also evaluated the membrane performance of pure water permeation (PWP) along with methyl blue and reactive red 120 dye rejection. M20-PPI1 with a weight percentage of PVP/Pluronic F108/PIA; 3%/1.5%/0.25%, exhibited the highest water flux of 12 L/m²h and dye rejection of 92% at a pH value of 8 for methyl blue dye removal whereas 82% at pH value of 2 for reactive red 120 dye removal. The findings proposed that M20-PPI1 is an effective pH-sensitive membrane for dye rejection.

Keywords: PH-Sensitive Membranes, Polymerized Itaconic Acid (PIA), Dye Removal, Pure Water Permeation (PWP)

1. Introduction

Membrane technology has been extensively applied in many separation industries due to its selectivity and permeability properties as one of the promising separation techniques [15]. Ladewig and Al-Shaeli [7] proposed polyethersulfone (PES) as an effective aromatic polymer in membrane fabrication including NF membranes which exhibited high thermal and mechanical resistance. Nevertheless, PES membranes had displayed hydrophobic properties that greatly affect the membrane flux and causes fouling, mainly in the accumulation and adsorption of substances on the pores and membranes surface. Thus, Safarpour *et al.*, [13] suggested that the modification of PES membranes either by chemical or physical means is vital, in order to enhance the hydrophilic properties of the fabricated membranes. Previous study that has been conducted by Milescu *et al.*, [9] reported that polyvinylpyrrolidone (PVP) is one of the various types of additives that had been included in the PES system, it acts as a pore former and structure controlling agents. Bolandi *et al.*, [1] also found that the employment of various surfactants as an additive in the dope solution had been successful in fabricating highly efficient membranes with high selectivity. Hamzah and co-workers [4,5] have employed pluronic F108, a non-ionic surfactant which led to the increase of water flux, membrane permeability and selectivity. Fan, Purkait and co-workers [3,11] also suggested that pH-

sensitive membranes exhibited remarkable improvements towards the permeability and separation performance of polymeric membranes. Besides, pH-sensitive membranes could also endow these membranes with pH-sensitive properties that make them adaptable to their application environment. The studies have stated that, functional groups such as carboxyl and pyridine displayed weak polyelectrolytes behaviors with a good pH-sensitive polymers. The utilization of polymerized itaconic acid, polyitaconic acid (PIA) into the membrane matrix fabrication would relatively produce this pH-sensitive properties. This is due to its highly functional polarity and polyelectrolyte properties of carboxylic acid groups (-COOH) and acetic acid groups (-CH₂COOH) in the PIA molecular structures.

Cationic and anionic dyes were the synthetic dyes that have been widely used in the dyeing process carrying hazardous and mutagenic properties. It is also hard to degrade where releasing them into the water bodies will result in adverse effects to health and environment including water pollution. Hence, Hosseinifard et al., [6] suggested that the development of an ideal and efficient material and method for dye removal of wastewater is greatly crucial. Mohammadnezhad et al., [10] stated that membrane technology had been widely recognized to resolve the environmental contamination issue. However, membrane filtration operations for wastewater treatment were prone to fouling that occurred mainly from the irreversible and reversible associations between the surface of membranes and numerous foulants. This happened mostly in hydrophobic membranes fabricated from various range of polymers.

The sole addition of additive was not sufficient in producing promising structural properties of the fabricated membranes by phase inversion approach. Therefore, the incorporation of other additives in the dope solution for membrane fabrication was required to overcome those limitations. No previous findings were found on the incorporation of polyitaconic acid (PIA) via blending into polymeric membrane matrix. Thus, the impact of the addition of PIA towards the PES matrix mainly in the membrane performance and structural properties were investigated. Besides, the novel PES membranes formulations were fabricated for the dye wastewater treatment application. The findings of this study have contributed new knowledge and data through incorporation of innovative materials in membrane fabrication which has showed its promising properties for the dye wastewater treatment. Hence, this study was effective as an effort of providing novel membrane for dye wastewater treatment in the removal of methyl blue (cationic) and reactive red 120 (anionic) dyes. In this study, PES membranes incorporated with polyvinylpyrrolidone, PVP (additive), pluronic F108 (non-ionic surfactant) and PIA of different concentrations, by dry/wet phase inversion were fabricated and characterized using appropriate instruments. Moreover, the PES membranes performance of the pure water permeation (PWP) was evaluated along with the methyl blue and reactive red 120 dyes rejection.

2. Methodology

2.1 Materials

Polyethersulfone (PES), n-Methyl-2-Pyrrolidone (NMP), polyvinylpyrrolidone (PVP) were purchased from Merck while pluronic F108, itaconic acid (IA), methyl blue dye and reactive red 120 dye were purchased from Sigma Aldrich. Ethanol and hexane were purchased from HmbG Chemicals.

2.2 Preparation of Polyitaconic Acid (PIA)

Itaconic acid (IA) need to undergo polymerization process to produce polyitaconic acid (PIA). Enchev [2] has proposed an effective procedure for the high conversion of itaconic acid

polymerization. Firstly, 100 g of IA was half neutralized with NaOH in 50 mL cold deionized water by cooling it with ice water, which then produced a concentrated solution. This concentrated solution was purged with nitrogen gas to create inert condition and was heated to 90 °C to 100 °C. Next, 70 wt% of tert-butyl hydroperoxide (tBHP) which act as an initiator was injected to initiate the polymerization process. The polymerized solution became viscous with the increase in time. After the solution turned into brittle solid, the polymerization process completed. Afterwards, the PIA samples were cooled down and were ground using mortar and pestle. The PIA samples were then stored in a closed container.

2.3 Fabrication of Membrane

The PES membranes were fabricated via phase inversion approach. Firstly, PES was added gradually into the round bottom flask, containing NMP at 50 °C to 60 °C of heating under constant stirring. This process was done in the fume hood, until homogenous polymer suspension was produced. PVP and pluronic were subsequently added to the polymer suspension after all of the PES granules have dissolved. Different wt% of PIA samples obtained were also incorporated together into the dope solution by blending. The dope solution was left stirring for 8 hours at 50 °C to 60 °C of heating. The formulation of dope solution according to each membrane samples was tabulated in table 1.

Table 1: Formulation of Dope Solution

Membrane	PES (wt%)	NMP (wt%)	H ₂ O (wt%)	PVP (wt%)	Pluronic (wt%)	PIA (wt%)
M20	20	75.00	5	3	1.5	0
M20-PPI1	20	70.25	5	3	1.5	0.25
M20-PPI2	20	70.00	5	3	1.5	0.50
M20-PPI3	20	69.75	5	3	1.5	0.75

The dope solution prepared was sonicated for 30 minutes to remove any air bubble presents. The dope solution was subsequently poured onto the glass plate and was casted slowly to the thickness of 100 µm to 150 µm by using casting knife. The glass plate was immersed in the coagulation bath containing distilled water (non-solvent) and later the membrane film was produced. The membrane film was stripped manually and carefully from the glass plate in the coagulation bath. After that, the membrane film was left immersed in the distilled water for 24 hours. Afterwards, the membrane film was immersed in ethanol for another 24 hours and in hexane for 3 hours consecutively. The PES membrane film was hung in the room temperature condition to let it dry.

2.4 Characterization and Performance Evaluation of Membrane

The membrane films were further characterized using UV-Vis, FTIR and TGA.

2.4.1 UV-Vis Spectrometer

UV-Vis spectrometer (Varian Cary 50) was operated at a wavelength range of 400 nm to 700 nm using single beam configuration to investigate the absorbance of stock and filtered methyl blue and reactive red 120 dye solutions. The standard calibration on the other hand was done by diluting the stock solution to the specific concentrations.

2.4.2 Fourier Transform Infrared (FTIR)

The membrane films were characterized using FTIR spectroscopy (Nicolet iS50) to give information on the functional groups present in the PES membranes. FTIR spectroscopy were

operated using attenuated total reflection (ATR) technique at wavenumber, λ ranging from 400 cm^{-1} to 4000 cm^{-1} .

2.4.3 Thermal Gravimetric Analyzer (TGA)

TGA (Mettler Toledo) was used to characterize the distribution of chemical components in the membrane formulation along with the thermal stability of the membrane samples based on the decomposition curve shift. This was done by placing about 15 mg samples into the platinum crucible. TGA was operated under nitrogen atmosphere with heating rate of 20 $^{\circ}\text{C}/\text{min}$ and flow rate of 50 mL/min at a temperature range of 30 $^{\circ}\text{C}$ to 1000 $^{\circ}\text{C}$.

2.4.4 Pure Water Permeation (PWP)

PWP was utilized to evaluate the pure water flux and permeability of every membrane film that were produced as stated by Lau and Ismail [8]. The dead-end filtration system was pressurized to the several pressure ranging from 3 to 7 bars containing membrane film with 30 mL distilled water in the pressurized cell. The water permeation was evaluated by quantifying the water passing through the membrane film or water flux, (J_w) at a rate of $\text{L}/\text{m}^2\text{h}$. The water permeation was computed using the Eq. (1). Where Q_p was the volume of permeate (L), A was the active surface area of membrane film (m^2), and Δt was the duration time of the permeation procedure (h).

$$J_w = \frac{Q_p}{A\Delta t}$$

(1)

2.4.5 Dye Rejection

The dye removal was also proceeded using the dead-end filtration system in pressurized condition ranging from 3 to 7 bars. The membrane was firstly soaked in distilled water of different pH values of 2, 4, 6 and 8 for 24 hours before testing. Afterwards, the dye rejection was tested with both methyl blue and reactive red 120 dye solution of matching pH values. This was based on the method that had been conducted by Fan *et al.*, [3]. The methyl blue and reactive red 120 dyes rejection percent were calculated using the Eq. (2). Where, C_p and C_f were denoted as the concentration of solute in permeate and feed consecutively.

$$R(\%) = 1 - \frac{C_p}{C_f} \times 100\%$$

(2)

3. Results

3.1 Fourier Transform Infrared (FTIR)

Figure 1 depicts the FTIR spectra of pure PES membrane, M20 and PES-PIA membranes incorporated with PIA, M20-PPI1, M20-PPI2 and M20-PPI3. M20-PPI1, M20-PPI2 and M20-PPI3 showed broad and strong peak ranging from 3403.74 cm^{-1} to 3431.32 cm^{-1} indicating O-H stretching vibration of itaconic acid. The intensity of O-H bands was slightly increased may due to the increment of PIA concentration incorporated in M20-PPI1, M20-PPI2 and M20-PPI3. All samples exhibited medium peak ranging from 3095.29 cm^{-1} to 3095.64 cm^{-1} specifying N-H stretching vibration of NMP. At the wavenumber ranging from 2871.70 cm^{-1} to 2873.28 cm^{-1} , both of the pure PES membrane and PES-PIA membranes depicted medium peak indicating C-H stretching vibration of Pluronic F108. Moreover, all samples showed medium to weak peak ranging from 1577.13 cm^{-1} to 1577.62 cm^{-1} specifying C=C stretching

vibration of phenyl structure of PES. The pure PES membrane and PES-PIA membranes also manifested C-H bending vibration of disubstituted, para phenyl structure of PES at the wavenumber ranging from 834.59 cm^{-1} to 870.78 cm^{-1} . Strong peak ranging from 698.87 cm^{-1} to 717.66 cm^{-1} was also depicted by all samples, M20, M20-PPI1, M20-PPI2 and M20-PPI3 indicating C-H bending vibration of disubstituted, ortho phenyl structure of NMP.

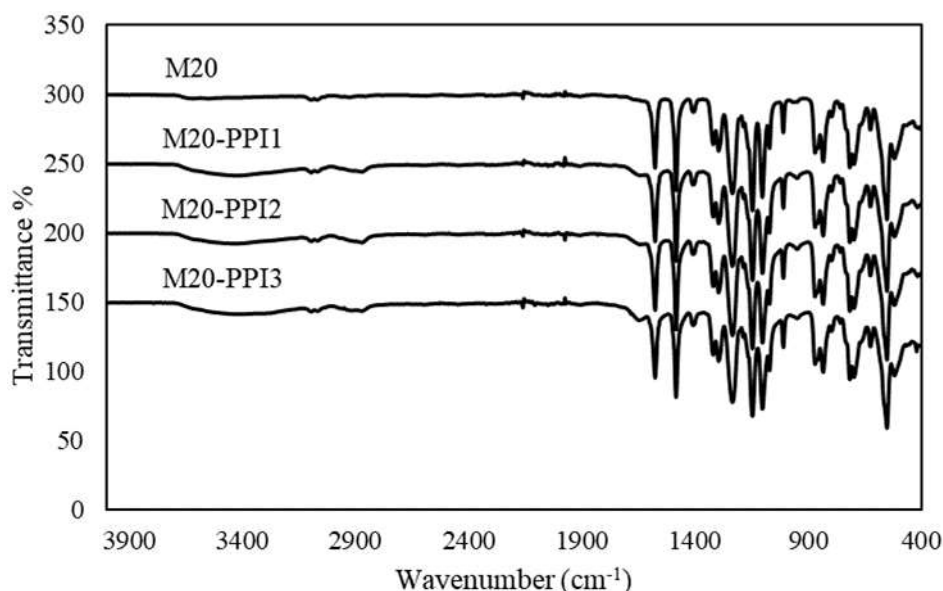


Figure 1: FTIR Spectra of Pure PES Membrane, M20 and PES-PIA Membranes Incorporated with PIA, M20-PPI1, M20-PPI2 and M20-PPI3

3.2 Thermal Gravimetric Analyzer (TGA)

Figure 2 illustrates the TGA thermogram of pure PES membrane, M20 and PES-PIA membranes incorporated with PIA, M20-PPI1, M20-PPI2 and M20-PPI3. The pure PES membrane displayed mainly two stages of TGA curve while the PES-PIA membranes samples depicted mainly three stages of TGA curve. The first stage decomposition started at the temperature of $40\text{ }^{\circ}\text{C}$ to $120\text{ }^{\circ}\text{C}$ with about 4% weight loss occurred at $110\text{ }^{\circ}\text{C}$ for M20-PPI1 and at $120\text{ }^{\circ}\text{C}$ for M20-PPI2 and M20-PPI3. This was due to the loss of water molecules and hydroxyl group attached to the PES-PIA membranes. The second stage took place at the temperature ranging from $400\text{ }^{\circ}\text{C}$ to $580\text{ }^{\circ}\text{C}$ with approximately 52% weight loss observed at $570\text{ }^{\circ}\text{C}$ for M20-PPI1, M20-PPI2 and M20-PPI3 and at $580\text{ }^{\circ}\text{C}$ for M20. This indicated the decomposition of polymer chain cleavage of the pure PES membrane and the PES-PIA membranes. The last stage was from $580\text{ }^{\circ}\text{C}$ to $1000\text{ }^{\circ}\text{C}$ with around 12% weight loss for all samples depicted the carbonization of the whole polymer chain. Basically, at the temperature of $570\text{ }^{\circ}\text{C}$ for M20-PPI1, M20-PPI2 and M20-PPI3 and $580\text{ }^{\circ}\text{C}$ for M20 were the main decomposition temperature. Although the thermal stability of PES-PIA membranes was slightly decreased from the pure PES membrane, they still displayed a high thermal stability with a constant decomposition temperature between the samples, M20-PPI1, M20-PPI2 and M20-PPI3. Basically, the incorporation of PIA has produced PES-PIA membranes with a considerably high thermal stability. The observations were supported by the findings of Sakthivel *et al.*, [14] and Rowley and Abu-Zahra [12] in their previous thermal characterization study.

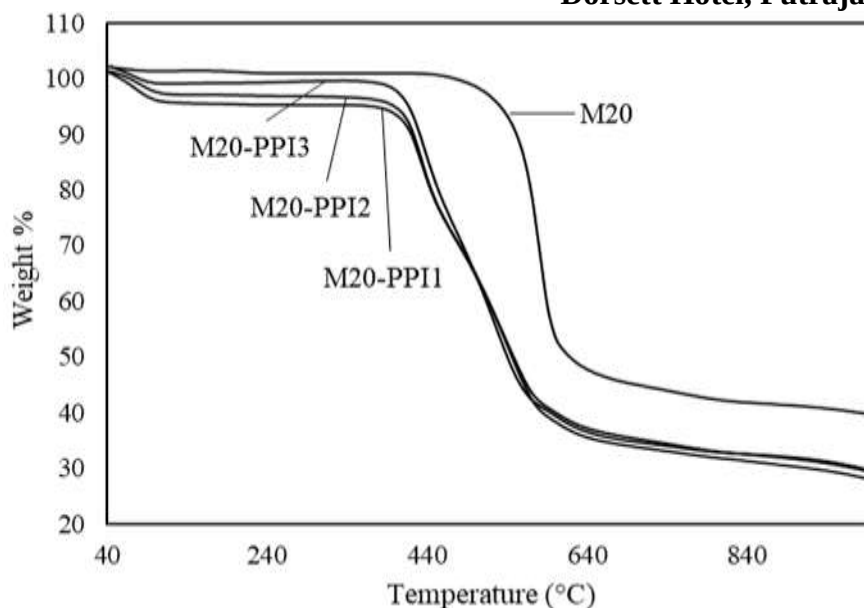


Figure 2: TGA Thermogram of Pure PES Membrane, M20 and PES-PIA Membranes Incorporated with PIA, M20-PPI1, M20-PPI2 and M20-PPI3

3.3 Pure Water Permeation (PWP)

Figure 3 displays the water flux of pure PES membrane, M20 and PES-PIA membranes incorporated with PIA, M20-PPI1, M20-PPI2 and M20-PPI3. From the graph, the pure PES membrane M20 showed incredibly low water flux with less than $0.1 \text{ L/m}^2\text{h}$. This was due to the hydrophobic properties of the PES polymeric membrane. M20-PPI1 on the other hand depicted the highest water flux with $11.73 \text{ L/m}^2\text{h}$ at the lowest incorporation of PIA which was 0.25%. It was found that the incorporation of small amount of PIA has produced the PES-PIA membranes with significantly high-water flux. This might be because of the largest pore size compared to the other PES-PIA samples which were M20-PPI2 and M20-PPI3. The incorporation of 0.5% PIA of M20-PPI2 have led to distinct decline of water flux to $0.42 \text{ L/m}^2\text{h}$. This was due to the reduction of the interspaces of pristine PES chains due to the incorporation of increasing PIA that led to the decrease in the size of pores. However, slight increase of water flux of M20-PPI3 with $0.51 \text{ L/m}^2\text{h}$ may due to the uneven pore distribution within the pore networks. It can be observed that the incorporation of only small amount of PIA for the PES-PIA membranes have improved the pure PES membrane permeability and hydrophilicity. The reasonings made were based on the previous findings studied by Fan et al., [3] in their pore size analysis investigation.

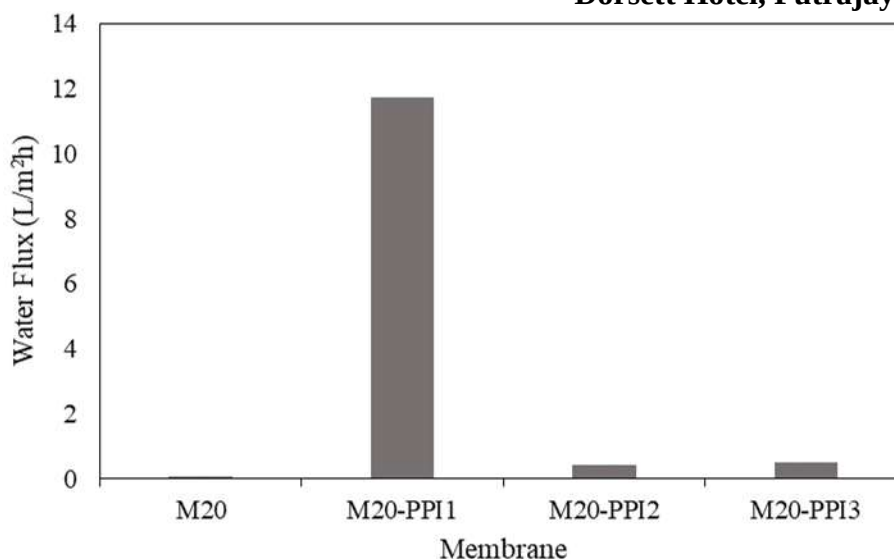


Figure 3: Water Flux of Pure PES Membrane, M20 and PES-PIA Membranes Incorporated with PIA, M20-PPI1, M20-PPI2 and M20-PPI3

3.4 Dye Rejection

Figure 4 shows the rejection % of methyl blue (MB) and reactive red 120 (RR120) of pure PES membrane, M20 and PES-PIA membranes incorporated with PIA, M20-PPI1, M20-PPI2 and M20-PPI3 at varying pH. From the graph, the pure PES membrane, M20 exhibited constant MB rejection % with 90% rejection. The high MB rejection % were due to the small pore size of pure PES membrane. However, it has very low water flux with very low permeability. The MB dye rejection % greatly increased along with the increment of pH for the blended PES-PIA membranes. The MB dye rejection % of M20-PPI1 at pH 2,4,6 and 8 were 18%, 19%, 90% and 92% respectively. M20-PPI2 MB dye rejection % were 13%, 53%, 87% and 89% at pH 2,4,6 and 8 accordingly. The MB dye rejection % at pH value 2,4,6 and 8 were 9%, 65%, 67% and 83% for M20-PPI3 subsequently. The findings of the study exhibited that the blended PES-PIA membranes showed the highest MB dye rejection at pH value of 8 with 92%, 89% and 83% for M20-PPI1, M20-PPI2 and M20-PPI3. On the other hand, the lowest dye rejection could be observed at pH value 2 with 9%, 13% and 18% for M20-PPI3, M20-PPI2 and M20-PPI1 MB dye removal. This was due to the swelling of PIA at high pH value that led to the decline in size of pores. The decline in pore size with the increment of pH have resulted in higher MB rejection % at higher pH. The inference made were based on the previous study conducted by Purkait et al., [11] in their research on blended pH-responsive membrane.

For the RR120 rejection, the pure PES membrane, M20 also displayed constant RR120 rejection % with 45% rejection. The RR120 dye rejection % declined along with the increment of pH for the blended PES-PIA membranes. The RR120 dye rejection % of M20-PPI1 at pH 2,4,6 and 8 were 82%, 48%, 42% and 1% respectively. M20-PPI2 RR120 dye rejection % were 35%, 17%, 6% and 3% at pH 2,4,6 and 8 accordingly. The RR120 dye rejection % at pH value 2,4,6 and 8 were 28%, 12%, 10% and 3% for M20-PPI3 subsequently. The findings of the study exhibited that the blended PES-PIA membranes showed the highest RR120 dye rejection at pH value of 2 with 82%, 35% and 28% for M20-PPI1, M20-PPI2 and M20-PPI3. On the other hand, the lowest dye rejection could be observed at pH value 8 with 1%, 3% and 3% for M20-PPI3, M20-PPI2 and M20-PPI1 RR120 dye removal. This has confirmed that the blended PES-PIA membranes were endowed with the pH-sensitive properties.

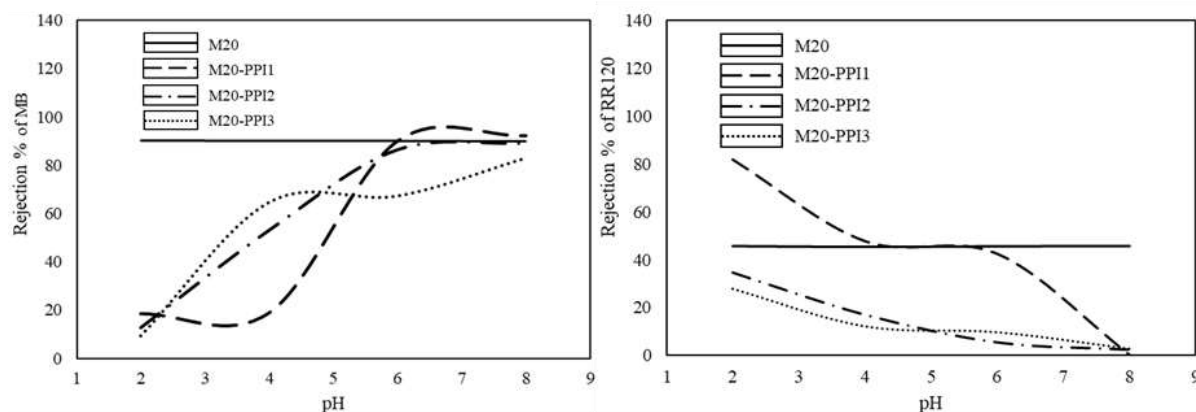


Figure 4: Rejection % of Methyl Blue (MB) and Reactive Red 120 (RR120) of Pure PES Membrane, M20 and PES-PIA Membranes Incorporated with PIA, M20-PPI1, M20-PPI2 and M20-PPI3 at varying pH.

4. Conclusions

From this study, PES membranes incorporated with PVP, pluronic F108 and PIA were successfully fabricated and characterized. Besides, the PES membranes performance of the pure water permeation (PWP) along with the methyl blue and reactive red 120 dyes rejection were also successfully evaluated. The study concluded that the FTIR analysis has confirmed the incorporation of PIA into the blended PES-PIA membranes. Besides, the TGA analysis has proved that the incorporation of PIA has produced PES-PIA membranes with a considerably high thermal stability. The study also showed that M20-PPI1 with a weight percentage of PVP/Pluronic F108/PIA; 3%/1.5%/0.25%, exhibited the highest water flux of 12 L/m²h and the highest dye rejection of 92% at a pH value of 8 for methyl blue dye removal whereas 82% at pH value of 2 for reactive red 120 dye removal. Moreover, the incorporation of PIA into the PES matrix has produced blended PES-PIA membranes with pH-sensitive properties.

Acknowledgement

This research was funded by a grant from Malaysian Nuclear Agency (USIM/MG/NUKLEAR/055012/70519).

References

- [1] Bolandi, S, F. Zokaee Ashtiani, A Okhovat, and M Boshrouyeh Ghandashtani. 2020. "The Effects of Surfactant Additives on Properties and Morphologies of Cellulose Acetate Membranes Prepared by Phase Inversion." *Theoretical Foundations of Chemical Engineering* 54 (5): 931–39.
<https://doi.org/10.1134/S0040579520050292>
- [2] Enchev, Dobromir D. 2017. "Synthesis and Properties of N-Alkyl(Phenyl)Amido-O-Methyl-1,2-Alkadienephosphonates." *Acta Scientifica Naturalis* 4 (1): 6–12.
<https://doi.org/10.1515/asn-2017-0002>
- [3] Fan, Kai, Guoqing Zhou, Jinjin Zhang, Haijun Yang, Jun Hu, and Zhengchi Hou. 2018. "PH-Sensitive Microfiltration Membrane Prepared from Polyethersulfone Grafted with Poly(Itaconic Acid) Synthesized by Simultaneous Irradiation in Homogeneous Phase." *Water Science and Technology* 78 (3): 602–10.
<https://doi.org/10.2166/wst.2018.330>
- [4] Hamzah, Norzakiah, Fasihah Johary, Rosiah Rohani, Syazrin Syima Sharifuddin, and Mohd Hafez Mohd Isa. 2020. "Development of Polyethersulfone (PES)/Reduced

- Graphene Oxide Nanocomposite Nanofiltration Membrane.” *Malaysian Journal of Fundamental and Applied Sciences* 16 (4): 418–21.
<https://doi.org/10.11113/mjfas.v16n4.1613>
- [5] Hamzah, Norzakiah, Rosiah Rohani, Abdul Rahman Hassan, Syazrin Syima Sharifuddin, and Mohd Hafez Mohd Isa. 2018. “Development of Chitosan/Pluronic F108/Polyethersulfone (PES) Nanofiltration (NF) Membrane for Oily Wastewater Treatment.” *AIP Conference Proceedings* 1972.
<https://doi.org/10.1063/1.5041235>
 - [6] Hosseiniard, Seyed Mehdi, Mohammad Ali Aroon, and Behnaz Dahrazma. 2020. “Application of PVDF/HDTMA-Modified Clinoptilolite Nanocomposite Membranes in Removal of Reactive Dye from Aqueous Solution.” *Separation and Purification Technology* 251 (April): 117294.
<https://doi.org/10.1016/j.seppur.2020.117294>
 - [7] Ladewig, Bradley, and Muayad Nadhim Zemam Al-Shaeli. 2017. *Fundamentals of Membrane Bioreactors*.
<https://doi.org/10.1007/978-981-10-2014-8>
 - [8] Lau, Woei Jye, and Ahmad Fauzi Ismail. 2017. *Nanofiltration Membranes*. Taylor & Francis Group. Vol. 3.
https://doi.org/10.1007/978-3-642-40872-4_2207-1
 - [9] Milesescu, Roxana A., C. Robert McElroy, Thomas J. Farmer, Paul M. Williams, Matthew J. Walters, and James H. Clark. 2019. “Fabrication of PES/PVP Water Filtration Membranes Using Cyrene®, a Safer Bio-Based Polar Aprotic Solvent.” *Advances in Polymer Technology* 2019.
<https://doi.org/10.1155/2019/9692859>
 - [10] Mohammadnezhad, Farrokh, Mostafa Feyzi, and Sirus Zinadini. 2019. “A Novel Ce-MOF/PES Mixed Matrix Membrane; Synthesis, Characterization and Antifouling Evaluation.” *Journal of Industrial and Engineering Chemistry* 71: 99–111.
<https://doi.org/10.1016/j.jiec.2018.09.032>
 - [11] Purkait, Mihir Kumar, Manish Kumar Sinha, Piyal Mondal, and Randeep Singh. 2018. “PH-Responsive Membranes.” In *Interface Science and Technology*, 25:39–66.
<https://doi.org/10.1016/B978-0-12-813961-5.00002-4>
 - [12] Rowley, Jiyoung, and Nidal H. Abu-Zahra. 2019. “Synthesis and Characterization of Polyethersulfone Membranes Impregnated with (3-Aminopropyltriethoxysilane) APTES-Fe₃O₄ Nanoparticles for As(V) Removal from Water.” *Journal of Environmental Chemical Engineering* 7 (1): 102875.
<https://doi.org/10.1016/j.jece.2018.102875>
 - [13] Safarpour, Mahdie, Vahid Vatanpour, and Alireza Khataee. 2016. “Preparation and Characterization of Graphene Oxide/TiO₂ Blended PES Nanofiltration Membrane with Improved Antifouling and Separation Performance.” *Desalination* 393: 65–78.
<https://doi.org/10.1016/j.desal.2015.07.003>
 - [14] Sakthivel, M., D. S. Franklin, and S. Guhanathan. 2016. “PH-Sensitive Itaconic Acid Based Polymeric Hydrogels for Dye Removal Applications.” *Ecotoxicology and Environmental Safety* 134: 427–32.
<https://doi.org/10.1016/j.ecoenv.2015.11.004>
 - [15] Zhao, Jing, Guangwei He, Guanhua Liu, Fusheng Pan, Hong Wu, Wanqin Jin, and Zhongyi Jiang. 2018. “Manipulation of Interactions at Membrane Interfaces for Energy and Environmental Applications.” *Progress in Polymer Science* 80: 125–52.
<https://doi.org/10.1016/j.progpolymsci.2017.12.002>

FABRICATION AND CHARACTERIZATION OF NANOCOILS CNT/PLA FILAMENT AT DIFFERENT FILLER'S PERCENTAGES AS ELECTROMAGNETIC WAVE ABSORBERS

Farah Hamizah binti Puad @ Paat^{1,2,*}
Nawrah Afnan binti Ashraf^{1,2}
Fadzidah binti Mohd Idris¹

¹ Kolej GENIUS Insan, Universiti Sains Islam Malaysia, Bandar Baru Nilai, 71800 Nilai, N. Sembilan.

² Faculty of Science, Universiti Sains Islam Malaysia, Bandar Baru Nilai, 71800 Nilai, N. Sembilan

Abstract: The rapid development of electronic technology and communication networks has resulted in high levels of electromagnetic interference (EMI) pollution, indirectly making electromagnetic wave (EMW) absorbers as of great interest among researchers. The characteristics of an ideal EMW absorber should be lightweight, thin, strong and capable of absorbing the EMW in a broad frequency range. Recently, carbon-based materials such as carbon nanotubes (CNTs) have gained good popularity for their unique properties including lightweight, high mechanical strength, good electrical and thermal conductivity and wide frequency bandwidth. However, carbon nanotubes are very conductive and tend to agglomerate due to the van der Waals forces. Therefore, the incorporation of CNTs into polymer helps to improve the EMW absorption mechanism. This research focuses on fabricating electromagnetic wave absorbers with nanocomposite based on CNT/PLA filament for additive manufacturing (3D printing). Polylactic acid (PLA) with different weight percentages of CNT fillers (1, 3, 5 wt%) was extruded to fabricate a composite filament of a diameter of 1.75 mm. The results show that different percentages of CNTs fillers have a significant effect on the morphological and crystallographic phase structure, Raman scattering and EMW absorption properties.

Keywords: EMI pollution, Carbon nanocomposites, fillers, reflection loss, 3D printing

1. Introduction

Electromagnetic interference (EMI) pollution has become a threat and severe issue to human health such as cancer which arises from the interference of various electromagnetic waves, either internal or external [1]. Prolonged exposure to EMI pollution not only affects the biological environment but also interferes with electronic performance resulting in signal disruption and dysfunction [2,3]. Therefore, the lightweight electromagnetic wave (EMW) absorber has aroused attention for limiting EMI pollution and protecting human as well as electronic systems [4,5].

In the last few years, electrically conductive metals are the most often used materials for EMW absorbers due to their strong conductivity [6]. However, the use of metal and metal composites is hindered by their limitations such as high density, corrosion and expensive [7,8]. To overcome this limitation, the advanced EMW absorber is now mainly focused on carbon nanocomposite material for its unique structure and properties. Carbon nanostructures such as

carbon nanotubes (CNTs), carbon black and graphene have gained popularity over metals because of excellent electrical conductivity, flexibility and corrosion resistance.

Among them, CNTs have great potential to be used as excellent EMW absorbers [9]. For instance, a 3D nitrogen-doped reduced graphene oxide/multi-walled carbon nanotube (NRGO/MWCNTs) composite aerogels by a facile hydrothermal and lyophilization two-step strategy was fabricated and reported by Shu *et al.*, [10]. They found that the composite aerogel with a filler content of 15 wt% displayed a minimum reflection loss of -53.3 dB at 6.5 GHz (C band) and broad absorption bandwidth of 5.2 GHz (from 11.1 to 16.3 GHz) under a small matching thickness of merely 2.0 mm. On the other hand, Chen *et al.*, [11] reported for the fabrication of MWCNTs/SiC nanocrystals/amorphous SiOC ceramic composites via pyrolyzed process, exhibited a great microwave-absorbing behaviour with the minimum reflection loss (RL) of -61.8 dB and effective absorption bandwidth (EBA) of 2.6 GHz (a thickness of 2.19 mm) at X-band.

Nevertheless, CNTs has a drawback in a massive production yield with low-cost material. The synthesis process is still restricted to carbon arc discharge and laser ablation which results in less production. Apart from increasing the quantity of CNTs production by using the chemical vapor deposition (CVD) process, incorporating CNTs with polymer is one of the promising ways to overcome this limitation which requires only less amount of CNTs as a filler in the composite [12,13]. According to Shahapurkar *et al.*, [14], polymer-based materials (such as polylactic acid (PLA), Acrylonitrile Butadiene Styrene (ABS), polycarbonate (PC), etc have the potential to attenuate the electromagnetic waves mainly by absorption, which has made them more popular in various applications, such as in civic and military technology [15,16]. PLA is a biodegradable polymer used as an alternative polymer resin for composite materials due to their remarkable features such as high tensile strength, sustainability, low melting point and adaptability in many applications [17,18].

Moreover, there are a lot of previous studies on the EM absorption properties of CNT/polymer that are being produced using various methods and sources, and according to Chen *et al.* [19], CNTs have a number of benefits, including a high dielectric loss, outstanding thermodynamic stability, and low-density EM. Even so, CNTs combined with a variety of materials demonstrated a better capacity to absorb EM waves than a CNTs alone.

The CNTs is tailor-made to maximize its absorption capacity of electromagnetic waves by integrating it in a polymer matrix as an efficient technique for the production of composite absorbers. An ideal EM wave absorber would have several benefits, such as high absorption efficiency, a wide absorption bandwidth, low density, and thin material thickness. However, the high material thickness and a limited effectiveness to absorb significant broadband have restricted other applications of CNTs polymer composites. Numerous significant works have already been completed to address these issues. Among these, one of the most efficient methods is to build a multi layered and porous structure. There is no question that a porous or hollow structure can reduce composite density while improving the absorption behaviour by raising the multi-reflecting and scattering percentage. The multi layered structure was shown to be effective in terms of achieving the impedance matching requirement [20].

1.1 EM Absorption Mechanism

EM materials have the ability to reflect, absorb, and transmit EMW when they are emitted. Note that the term "absorption" mostly relates to the ability of EM energy to be converted to

heat via intrinsic magnetic or dielectric loss. In general, the mechanisms of EMW absorption of CNTs-based composites are related to impedance matching and attenuation principles [19].

EMW absorption occurs when most of the EMW signal must be penetrating inside the absorbing material. When the incident EMW reaches the surface of the loss materials, the EMW entering the absorbers must be attenuated as much as possible. Meanwhile, the absorptivity of absorbers are related to the parameters of the relative complex permittivity ($\epsilon_r = \epsilon' - i\epsilon''$) and the relative complex permeability ($\mu_r = \mu' - j\mu''$). Here, the real part (ϵ' and μ') relates to the energy storage capacity, while the imaginary part (ϵ'' and μ'') is related to the energy dissipation. As per transmission line theory, the reflection loss (RL), can be calculated using the following Eq. (1) and Eq. (2) which contribute to almost half of the absorption effect [21]:

$$Z_{in} = Z_o \sqrt{\mu_r / \epsilon_r} \tanh[j(2\pi f d / c \sqrt{\epsilon_r \mu_r})] \quad (1)$$

$$RL (dB) = 20 \log \left| \frac{(Z_{in} - Z_o)}{(Z_{in} + Z_o)} \right| \quad (2)$$

The input impedance in this case is Z_{in} and Z_o is impedance of free space. While, ϵ_r and μ_r are relative complex permittivity and relative complex permeability respectively. On the other hand, f stands for the incident wave frequency, c for the speed of light and, d for the thickness of EMW absorption sample.

From theoretical stand point, an ideal impedance match can be obtained when the values of ϵ_r and μ_r are very close, that is, the value of $|Z_{in} / Z_o|$ near to 1. This way, the incident wave passes completely through the absorbers without reflection which resulting in an ideal reflection loss. Thus, incident EMW permeates into absorbers completely without reflection, and the corresponding reflection loss will be optimal [22].

Once an incident EMW penetrated the lossy material, its ability to attenuate the EMW became a critical indicator of how well it could absorb the energy. The following equation can be used to obtain the attenuation capability, which is called the attenuation constant (α) and can describe the intensity of absorption in Eq. (3) [23]:

$$\alpha = \frac{\sqrt{2}}{c} \pi f \sqrt{(\epsilon''\mu'' - \epsilon'\mu') + \sqrt{(\epsilon''\mu'' - \epsilon'\mu')^2 + (\epsilon'\mu'' + \epsilon''\mu')^2}} \quad (3)$$

Where, ϵ' and μ' are respectively denote as the real permittivity and real permeability. Whereas, ϵ'' and μ'' represent the imaginary part for both permittivity and permeability respectively.

In addition, the ratio between the imaginary and real parts of permittivity, or the dielectric loss tangent ($\tan \delta_\epsilon$) based on the following equation, Eq. (4) can be used to determine the dielectric loss [24]:

$$\tan \delta_\epsilon = \frac{\epsilon''}{\epsilon'} \quad (4)$$

Dielectric loss basically refers to the attenuation of EMW that result from contact with absorbers. In other words, absorbers respond to an applied oscillating EM field to convert energy from their electric field form to their thermal form. The main components of dielectric loss are conductivity loss and polarisation loss. Polarisation loss is primarily caused by ionic, electronic, interfacial, and dipole orientation polarisation. The numerous surface atoms, unsaturated bonds, and dangling bonds of CNTs used as fillers in EM absorbers, exhibit

interfacial polarisation and multiple EMW scattering. These are typical techniques for EM attenuation in composite absorbers and are crucial to the EM absorption mechanism of CNTs [19].

2. Methodology

2.1 Material Characteristics

PLA pellets with 98 % of L-lactide (trademark Luminy® LX575), were purchased from Total Corbion PLA (Thailand). The density and melt flow index was 1.24 g/cm³ and 7 g/10 min (210 °C/2.16 kg). This grade has a glass transition temperature (T_g) of 60 °C and a melting temperature (T_m) of 165 °C. PLA pellets were dried in the oven at 40 °C for 3 hours to remove moisture. In addition, synthesized CNTs produced by chemical vapor deposition (CVD) method with an average diameter size of 163.1 nm were used.

The synthesized CNTs were characterized by Field Emission Scanning Electron Microscope (FESEM, JSM-IT800 Schottky), X-Ray Diffraction (XRD, Rigaku MiniFlex 600), High-Resolution Transmission Electron Microscope (HRTEM, FEI Tecnai G2 20 S-TWIN), Energy-Dispersive X-Ray Spectroscopy (EDX), and Raman Spectroscopy (Horiba XploRA PLUS). The reflection loss (RL) value of the 3D printed EMW absorber was observed via PNA Vector Network Analyzer (VNA).

2.2 Sample Preparation of CNT

In this study, CNTs were synthesized from a mill scale waste via CVD method which mill scale as a catalyst; ethanol as a carbon source and argon gas as a carrier gas. The mill scale was first crushed for 8 hours with high energy ball milling (HEBM) to obtain smaller starting particle size. 0.1 g of mill scale was loaded into the alumina boat and placed in the furnace for CNTs growth. The temperature of the furnace was raised from room temperature to 700 °C at a rate of 5 °C/min within 2 hour 20 minutes. During the reaction, ethanol was introduced for 30 minutes while the deposition temperature was maintained. Finally, the furnace was let to cool down before the whole system was turned off. The resulting synthesized CNTs were kept in a vial for further measurement and characterization.

2.3 Development of CNT/PLA Filament for 3D Printing

The filaments were prepared with Siemens Twin Screw Extruder (ratio length, L over diameter, D, L/D= 24). At first, CNTs with different weight percentages (0, 1, 3, 5 wt%) were added into PLA polymer and manually mixed as shown in Table 1. 500 g of CNT/PLA mixture was inserted into the feeder gradually to avoid clumping during the melting process. There were three heating zones correspond to the feed zone, melting zone and melt processing zone and the temperature was fixed around 180-190 °C respectively. Meanwhile, the screw speed was set at 10-12 rpm. At the die, the extruded strand of the CNT/PLA composite was cooled down in a water bath and passed through a spooler to produce a uniform diameter size of the filament wire (1.75 ±0.1 mm).

Table 1: Percentage weight of CNTs

Weight % of CNT 8h	Mass CNT/PLA (g)
0%	500
1%	
3%	
5%	

After the extrusion stage, the filament was printed by using FlashForge Creator Pro 2 FDM 3D printer. The 3D model of a flat square with various thicknesses (1, 2, 3 mm) was designed using TinkerCAD software and exported as STL file format. The dimension of the 3D model refers to the waveguide size as shown in Table 2. The filament was fed into a 0.4 mm diameter nozzle. The parameter of the FDM 3D printing process was set as follows: the nozzle temperature (210 °C), bed temperature (40 °C), layer height (0.18 mm), infill density (100 %), printing speed (60 mm/s); these variables were kept constant for all model.

Table 2: Dimension of 3D printed model

Waveguide name (EIA standard)	Frequency band	Inner dimensions of waveguide opening (mm)	Thickness of 3D model (mm)
WR112	7.05 to 10 GHz	28.4988 x 12.6238	1,2,3
WR75	10.00 to 15 GHz	19.05 x 9.525	1,2,3

2.4 Materials' Characterization

The X-ray diffraction (XRD) was recorded on an X-ray diffraction (XRD, Rigaku MiniFlex 600) instrument with Cu K radiation operated at maximum power 600 W (40 kV- 15 mA). The morphology and microstructures were characterized by a field emission scanning electron microscope (FESEM, JSM-IT800 Schottky) and a high-resolution transmission electron microscope (HRTEM, FEI Tecnai G2 20 S-TWIN). The Raman scattering was characterized by using Raman Spectroscopy (Horiba XploRA PLUS). For the EM measurement, a vector network analyzer (VNA, Agilent PNA N5227A) was used to measure the scattering parameters (S-parameters) in 10 MHz-67 GHz. Figure 1 shows the whole process scheme of 3D printed model.

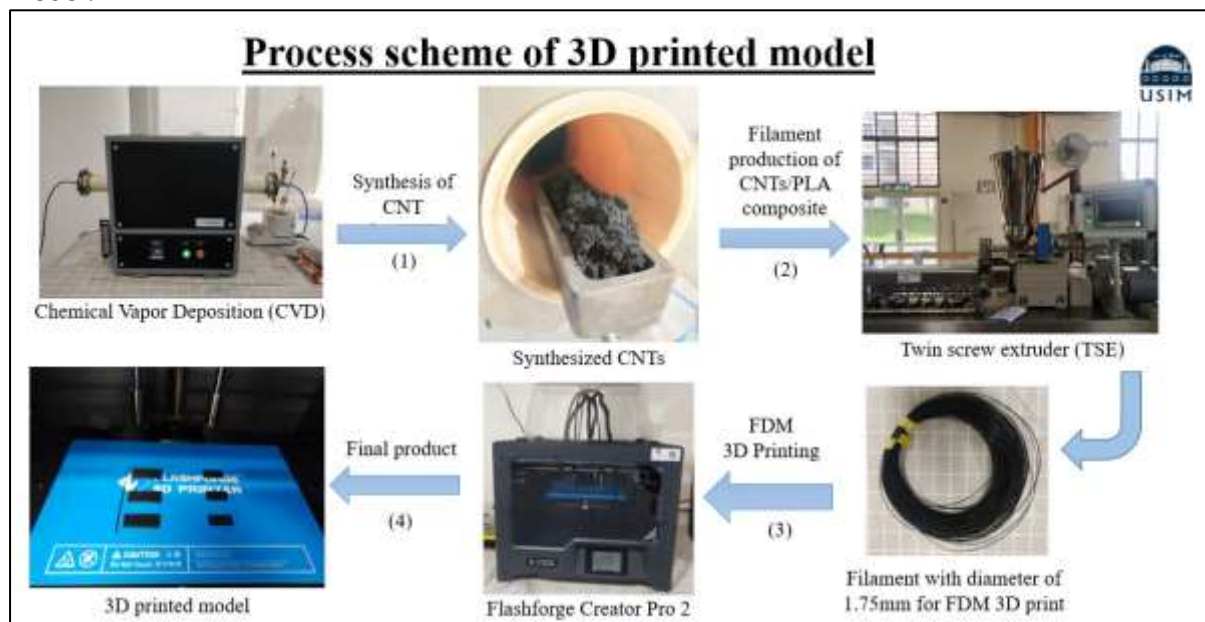


Figure 1: Process Scheme of 3D Printed Model.

3. Results and Discussion

3.1 Morphological and particle size of synthesized CNTs

The surface morphology of the mill scale was analyzed based on FESEM results. Figure 2 shows the image of raw mill scale waste and Figure 3 shows the mill scale milled for 8 hours. They present a slightly similar morphology but differ in terms of particle size. The average particle size of mill scale milled at 8 hours is 278.4 nm meanwhile the particle size of raw mill scale waste is 415.5 nm respectively. The milling time is believed to be the cause of the size reduction [25]. On the other hand, Figure 4 shows and revealed that the CNTs are highly entangled with each other. The average diameter of synthesized CNTs 8h is 163.1nm. The structure of CNTs was typically straight, twisted and have a high tendency to form a bundle and entanglement due to the Van der Waals interaction. Furthermore, the average diameter of CNTs is decreased compared to the average particle size of mill scale milled at 8 hours. This commonly happens after the production giving the lower volume fraction of the CNTs.

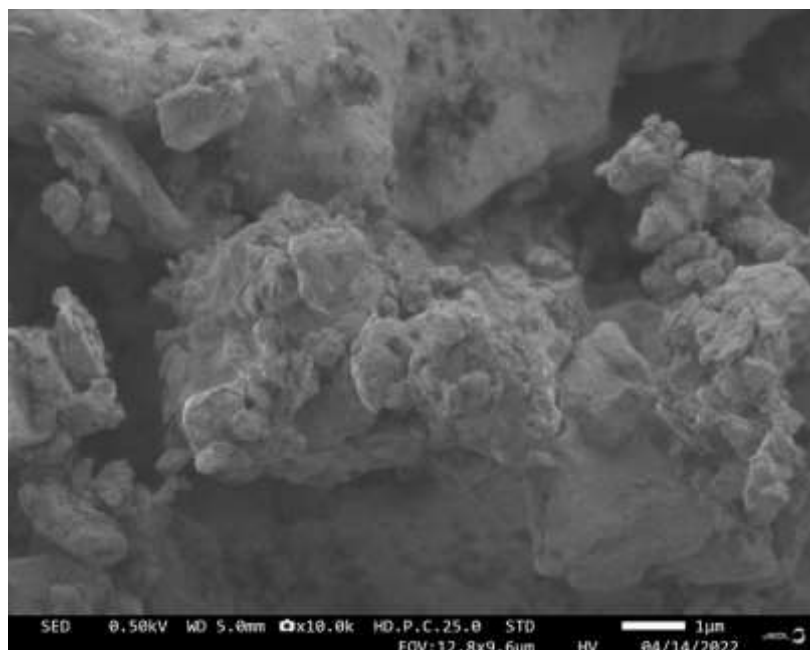


Figure 2: FESEM Image of Raw Mill Scale Waste

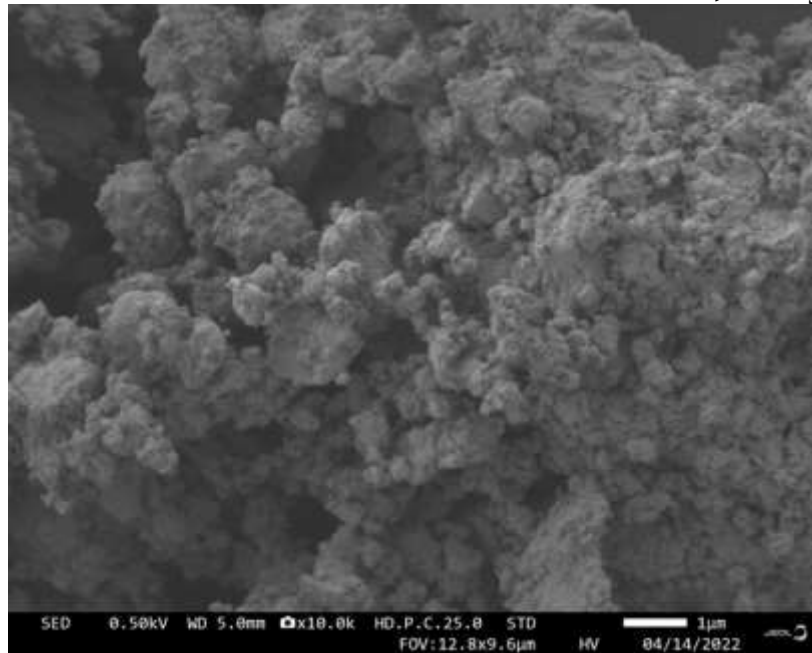


Figure 3: FESEM Image of Mill Scale Milled For 8 Hours

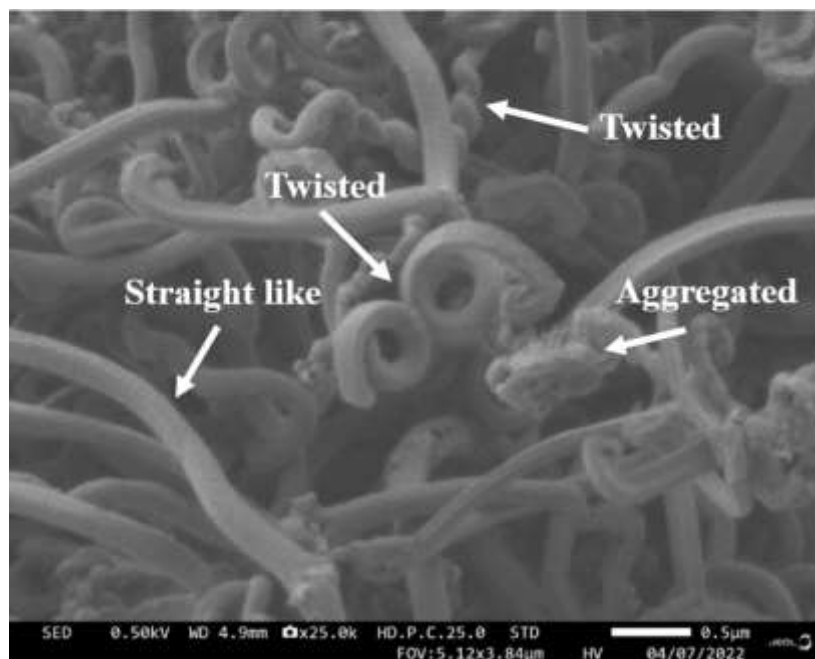


Figure 4: FESEM Image of Synthesized Cnts (Mill Scale Milled For 8 Hours)

3.2 X-ray Diffraction Analysis

The crystallinity and phase analysis of CNTs and mill scale was observed by XRD as shown in Figure 5.

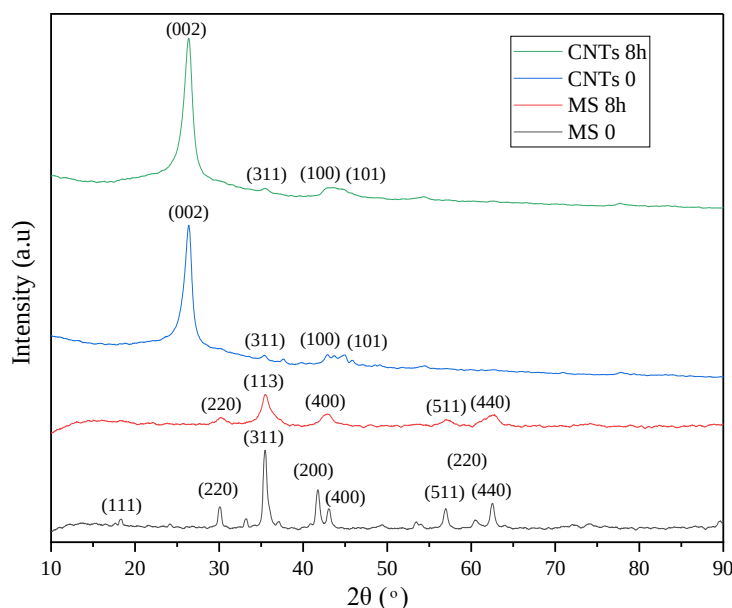


Figure 5: XRD Spectra of Raw Mill Scale (MS0), Mill Scale Milled At 8 Hours (MS 8h), Synthesized Cnts 0 And Synthesized Cnts 8h.

In general, mill scale is comprised primarily of wustite (FeO) and magnetite (Fe_3O_4). The results from XRD confirmed that the main phase for the raw mill scale is magnetite (Fe_3O_4) and wustite (FeO). However, for the sample mill scale milled at 8 hours, wustite peaks disappeared and showed only the presence of magnetite. As seen from the patterns, the XRD peaks for the sample raw mill scale corresponded to the diffraction pattern characteristic of the magnetite structure with (JCPD: 00-019-0629) observed at the $2\theta(^{\circ}) = 18.270^{\circ}$, 30.095° , 35.423° , 43.053° , 56.944° , 62.516° respectively, belonging to (111), (220), (311), (400), (511), (440) and wustite at $2\theta(^{\circ}) = 41.928^{\circ}$ (200) and $2\theta(^{\circ}) = 60.766^{\circ}$ (220) with reference code (JCPD: 00-006-0615). This formation of iron oxide on the surface of the mill scale commonly occurs during the hot rolling process involving a high temperature [26]. The intense peak at $2\theta(^{\circ}) = 26.382^{\circ}$ of CNTs represents the characteristic graphitic peak formed by the tubular structure of the carbon atoms in the sample with (002) planes with the reference code data (JCPD: 00-041-1487). Considering the spectral profile of CNTs, the weak broad XRD peak at $2\theta = 42.402^{\circ}$ and $2\theta = 44.600^{\circ}$ with (JCPD: 00-023-0064) is assigned to graphitic (100) and (101) crystalline lattice that related with Bragg law. In addition, the presence of magnetite (Fe_3O_4) can be seen at $2\theta = 35.423^{\circ}$ in the plane (311) with the reference code (JCPD: 00-019-0629). The (002) peak of CNTs 8h and CNTs 0 shows an increase FWHM pattern intermediate between amorphous carbon and crystalline due to the smaller size particle of CNTs. It is reported that the smaller the size of particle the greater the value of FWHM (Singh et al., 2010). The peak shift will broaden as the FWHM value increase and this properties indicate to have higher EMW absorption than crystalline CNTs [27].

3.3 Elemental Analysis

The elemental composition of synthesized CNTs of mill scale milled for 8h was detected with an Energy-Dispersive X-ray (EDX). Figure 6 illustrates the presence of oxygen (O), copper (Cu), iron (Fe) and carbon (C) peaks. The appearance of a high carbon peak is from the production of CNTs and the presence of iron (Fe) due to the use of mill scale as a catalyst. Meanwhile, the existence of the element copper (Cu) may be due to the sample preparation for the characterization involving a copper grid.

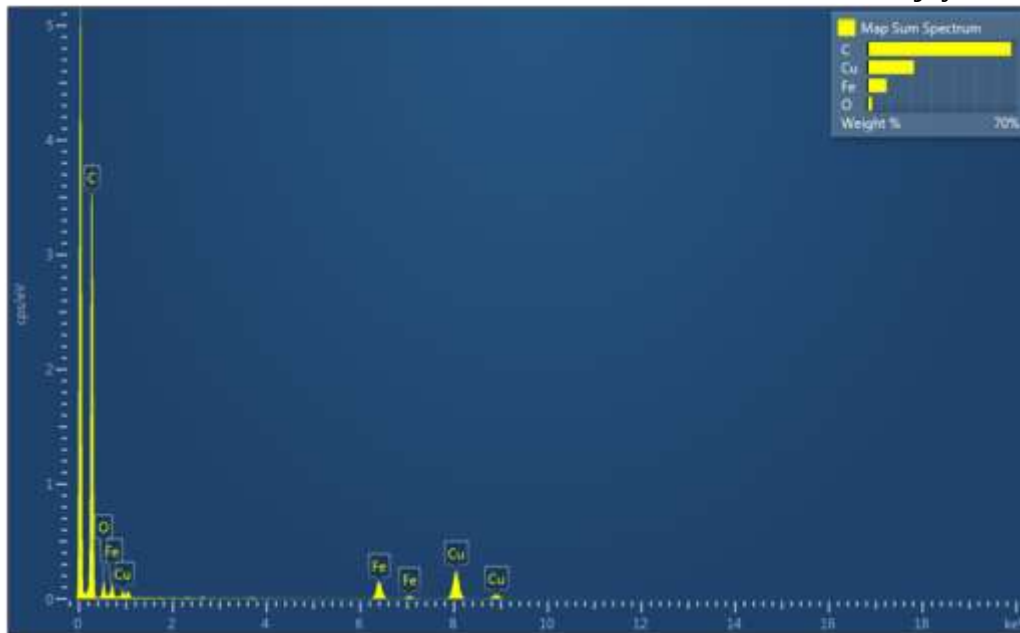


Figure 6: EDX Pattern of Synthesized Cnts 8h

3.4 HRTEM

High-resolution TEM (HRTEM) image of synthesized CNTs 8h exhibits layers of conductive tube walls in a uniform arrangement as shown in Figure 7(a). The average external diameter of CNTs 8h is 90.027 nm while the internal diameter is 39.074 nm. Figure 7(b) shows the coaxial cylindrical layers with a d-spacing of 1.4Å. The wall thickness of CNTs is 25.4765 nm which produces 176.9201 layers of carbon nanotubes.

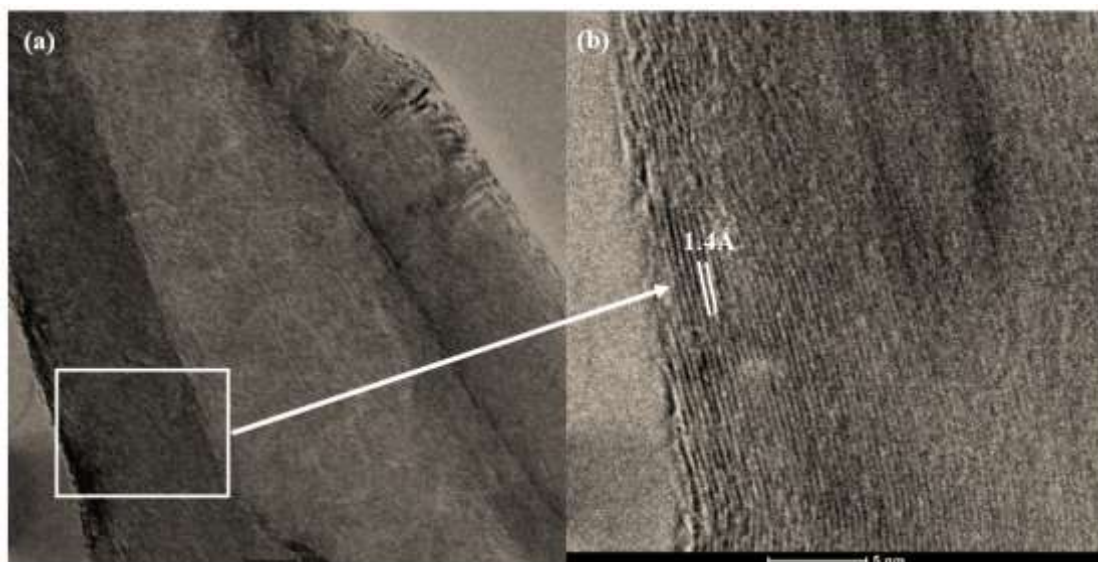


Figure 7: a) HRTEM Image of Synthesized Cnts 8h B) Enlarge Image of The Tubular Structure of Cnts 8h

According to [26], MWCNT comprises two or more concentric cylindrical shells that coaxially form around a central hollow core due to the van der Waals forces, with an average spacing of 0.34 nm, similar to graphite. MWCNTs have an external diameter of 5-50 nm and a length of 100 nm to several centimetres. The multi-layer structure of CNTs indicates the excellent absorption of EMW due to the multiple reflections between the layered interfaces [27].

3.5 RAMAN Scattering

The vibrational mode and Raman spectra of synthesized CNTs 8h are illustrated in Figure 8. The Raman spectrum has two major peaks known as D-band and G-band. The D-band represents a defect in the graphitic structure that is caused by out of plane vibration form at around 1335 cm^{-1} , whereas the G-band is caused by in plane vibration of the C-C bond from the sp^2 orbital indicate graphitic structure form at around 1570 cm^{-1} [28]. The ratio of I_D/I_G can estimate the degree of graphitization and defect density.

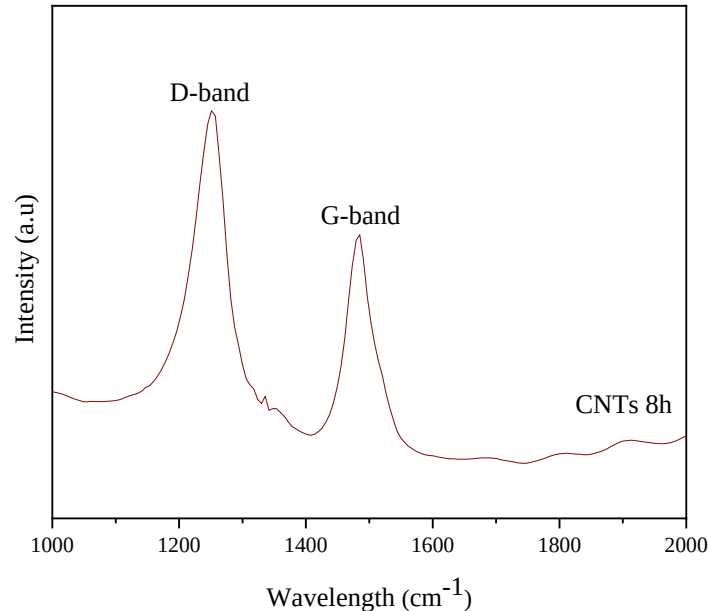


Figure 8: Raman Spectrum of Synthesized Cnts Of Mill Scale Milled At 8 Hours

Based on Figure 8, the result of the intensity ratio of I_D/I_G is 1.3010 which suggested that the CNTs have a high defect density and lower graphitization. The presence of defects in the structure is desirable because it shows a lower conductivity helping in an EMW absorption performance. Thus, the higher D-band also shows that it was related with the production of more multi-walled CNTs compared to single-walled CNTs.

3.6 EM Absorption Performance

This study measures the EMW absorption coefficient in three sets of samples of different CNTs thicknesses. Each thickness of CNTs has different filler's percentages of 0%, 1%, 3% and 5% filler percentages of CNTs for the frequency from 10 GHz to 15 GHz as shown in Figure 9.

As microwave energy penetrates a lossy dispersive material which in this case CNT/PLA composite, EM field interactions with the substance's molecular and electronic structures result in heating of the material. In addition, microwave radiation's energy is partially reflected, absorbed, and transmitted when it penetrate a solid surface. Transmittance (T), reflectance (R), and absorbance (A) can be computed using the formula in Eq. 5-7, according to the analysis of S-parameters [32] :

$$T = |S_{12}|^2 = |S_{21}|^2 \quad (5)$$

$$R = |S_{11}|^2 = |S_{22}|^2 \quad (6)$$

$$A = 1 - R - T \quad (7)$$

Where the scattering parameters of the two port network systems, S11 or S22 and S12 or S21, represent the reflection loss and transmission.

Based on Figure 9(a)-(c), there are fluctuations in each of the absorption coefficient graph with several high peaks and low peaks. In Figure 9(b), there are four obvious high peaks for samples with 2 mm thickness. The highest absorption coefficient peaks for all CNTs percentages occurred at the frequency range of 7 GHz to 8 GHz with the absorption coefficient value of 0.81 (CNT 0%), 0.77 (CNT 1%), 0.79 (CNT 3%) and 0.8 (CNT 5%). Another high peaks occurred at the frequency between 9 GHz to 10 GHz with the coefficient value of 0.72 (CNT 0%), (CNT 1%), (CNT 3%) and (CNT 5%). While the other two high peaks occurred at 12 GHz to 13 GHz and 13 GHz to 15 GHz. There is also lowest peak of absorption coefficient at the frequency range from 8 GHz to 9 GHz with CNTs 1% is the has the lowest absorption coefficient value which is 0.62.

The highest absorption peak occurred in the frequency range 7-8 GHz for all thickness and CNTs percentages. Meanwhile, the lowest absorption coefficient peak for all thickness and CNTs percentages happened at the frequency range of 8 GHz to 9 GHz except for sample CNTs 1% with 1 mm thickness, where the lowest peak occur at 10.3 GHz with absorption coefficient value of 0.6.

From Figure 9(a)-(c), the overall peak shift trend between the thickness of sample and the absorption coefficient peaks. The peaks absorption coefficient have a slight change of peaks position to the left as the thickness increases. For an instance, in Figure 9(c)-(c), the highest absorption peak for CNTs 1% of 1 mm is 0.84 at frequency 8GHz. Meanwhile, for CNTs 1% with 2 mm and 3 mm thicknesses the highest absorption coefficient peak is at 0.77 and 0.88 at frequency 7.14 GHz and 7.29 respectively.

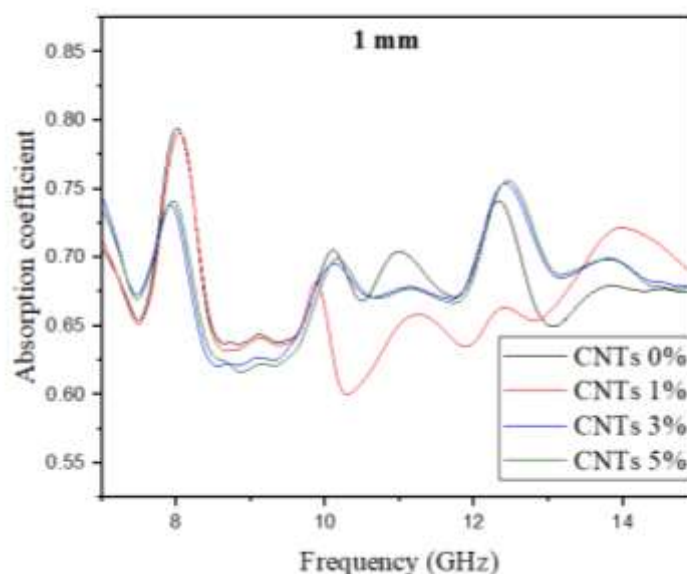


Figure 9 (A): Absorbance Coefficient Of 1mm Cnts

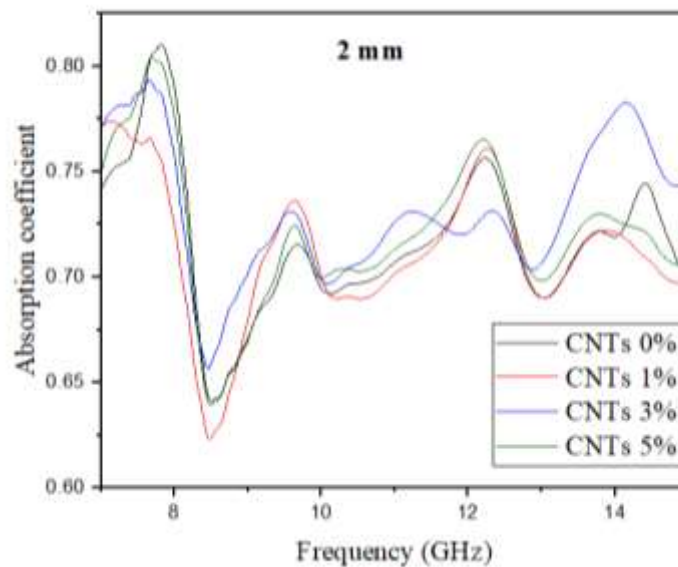


Figure 9 (B): Absorbance Coefficient Of 2 Mm Cnts

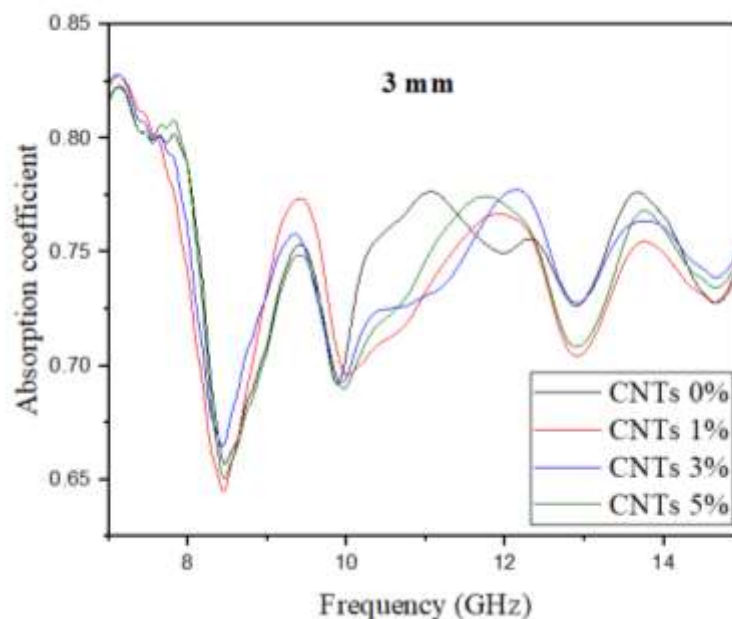


Figure 9 (C): Absorbance Coefficient Of 3 Mm Cnts

Figure 9: Absorbance coefficient graphs of three different thicknesses of CNTs; (a) CNTs 1 mm, (b) CNTs 2 mm, and (c) CNTs 3 mm with four different filler percentages

4. Conclusions

In summary, a lightweight, thin, strong absorption, wideband and low cost EMW absorber was developed. 3D printed CNTs/PLA with a different weight percentage on a thickness of 1,2,3 mm were fabricated by using the Fused deposition modelling (FDM) method. The absorption coefficient peak from the VNA result shows the fluctuation pattern in which the highest peak varies depending on the frequency. For example, CNT 1% with 2 mm thickness has the highest

absorption coefficient peaks for all CNTs percentages occurred at the frequency range of 7 GHz to 8 GHz with the absorption coefficient value of 0.81 (CNT 0%), 0.77 (CNT 1%), 0.79 (CNT 3%) and 0.8 (CNT 5%). Another high peaks occurred at the frequency between 9 GHz to 10 GHz with the coefficient value of 0.72 (CNT 0%), (CNT 1%), (CNT 3%) and (CNT 5%). Also, as the thickness increases, the absorption coefficient peak shifts toward the lower frequency. The absorbing mechanism of CNTs/PLA is related to the weight percent and size of the particles of CNTs. The morphology, phase analysis, particle size, and Raman spectrum of synthesized of CNTs were characterized via FESEM, XRD, EDX, HRTEM, RAMAN, and VNA. The small particle size of the mill scale milled at 8 hours (278.4 nm) contributed to the greater value of FWHM indicated to have a high EMW absorption.

Acknowledgement

The authors acknowledge the Ministry of Higher Education Malaysia for financial support through Fundamental Research Grant Scheme (FRGS/1/2020/STG05/ USIM/02/3) (USIM/FRGS/KGI/KPT/52020) and Universiti Sains Islam Malaysia for financial support under USIM-RACER Grant (PPPI/USIMRACER_0120/KGI/051000/11920).

References

- [1] Batool, S, A Bibi, F Frezza, and F Mangini. n.d. "Benefits and Hazards of Electromagnetic Waves, Telecommunication, Physical and Biomedical: A Review."
- [2] Hlo, Klaudia. 2020. "Nanoscale Advances Progress in Polymers and Polymer Composites Used as efficient Materials for EMI Shielding." <https://doi.org/10.1039/d0na00760a>.
- [3] Tian, Ke, Danrong Hu, Quan Wei, Qiang Fu, and Hua Deng. 2023. "Recent Progress on Multifunctional Electromagnetic Interference Shielding Polymer Composites." *Journal of Materials Science and Technology* 134 (January): 106–31. <https://doi.org/10.1016/j.jmst.2022.06.031>.
- [4] Liang, Xiaohui, Bin Quan, Yansheng Sun, Guangbin Ji, Yanan Zhang, Jianna Ma, Daoran Li, Baoshan Zhang, and Youwei Du. 2017. "Multiple Interfaces Structure Derived from Metal–Organic Frameworks for Excellent Electromagnetic Wave Absorption." *Particle and Particle Systems Characterization* 34 (5). <https://doi.org/10.1002/ppsc.201700006>.
- [5] Savi, Patrizia, Mauro Giorcelli, and Simone Quaranta. 2019. "Multi-Walled Carbon Nanotubes Composites for Microwave Absorbing Applications." *Applied Sciences (Switzerland)* 9 (5). <https://doi.org/10.3390/app9050851>.
- [6] Zhang, Zhiwei, Zhihao Cai, Ziyuan Wang, Yaling Peng, Lun Xia, Suping Ma, Zhanzhao Yin, and Yi Huang. 2021. "A Review on Metal–Organic Framework-Derived Porous Carbon-Based Novel Microwave Absorption Materials." *Nano-Micro Letters*. Springer Science and Business Media B.V. <https://doi.org/10.1007/s40820-020-00582-3>.
- [7] Lee, Seung Hwan, Yeongbeom Lee, Myung Geun Jang, Chonghun Han, and Woo Nyon Kim. 2017. "Comparative Study of EMI Shielding Effectiveness for Carbon Fiber Pultruded Polypropylene/Poly(Lactic Acid)/Multiwall CNT Composites Prepared by Injection Molding versus Screw Extrusion." *Journal of Applied Polymer Science* 134 (34): 1–11. <https://doi.org/10.1002/app.45222>.
- [8] Jiang, Dawei, Vignesh Murugadoss, Ying Wang, Jing Lin, Tao Ding, Qian Shao, Chao Wang, et al. 2019. "Electromagnetic Interference Shielding Polymers and Nanocomposites - A Review Electromagnetic Interference Shielding Polymers And" 3724. <https://doi.org/10.1080/15583724.2018.1546737>.
- [9] Li, Qi, Zheng Zhang, Luping Qi, Qingliang Liao, Zhuo Kang, and Yue Zhang. 2019. "Toward the Application of High Frequency Electromagnetic Wave Absorption by Carbon

Nanostructures.” *Advanced Science*. John Wiley and Sons Inc. <https://doi.org/10.1002/advs.201801057>.

[10] Shu, Ruiwen, Zongli Wan, Jiabin Zhang, Yue Wu, and Jianjun Shi. 2021. “Synergistically Assembled Nitrogen-Doped Reduced Graphene Oxide/Multi-Walled Carbon Nanotubes Composite Aerogels with Superior Electromagnetic Wave Absorption Performance.” *Composites Science and Technology* 210 (July). <https://doi.org/10.1016/j.compscitech.2021.108818>.

[11] Chen, Lixin, Jia Zhao, Lei Wang, Fei Peng, Hu Liu, Jiaoxia Zhang, Junwei Gu, and Zhanhu Guo. 2019. “In-Situ Pyrolyzed Polymethylsilsesquioxane Multi-Walled Carbon Nanotubes Derived Ceramic Nanocomposites for Electromagnetic Wave Absorption.” *Ceramics International* 45 (9): 11756–64. <https://doi.org/10.1016/j.ceramint.2019.03.052>.

[12] Munir, Arshad. 2017. “Microwave Radar Absorbing Properties of Multiwalled Carbon Nanotubes Polymer Composites: A Review.” *Advances in Polymer Technology*. John Wiley and Sons Inc. <https://doi.org/10.1002/adv.21617>.

[13] Barathi Dassan, Emayaruba G., Aslina Anjang Ab Rahman, Mohd Shukur Zainol Abidin, and Hazizan Md Akil. 2020. “Carbon Nanotube-Reinforced Polymer Composite for Electromagnetic Interference Application: A Review.” *Nanotechnology Reviews*. De Gruyter Open Ltd. <https://doi.org/10.1515/ntrev-2020-0064>.

[14] Shahapurkar, Kiran, Mengistu Gelaw, Vineet Tirth, Manzoore Elahi M. Soudagar, Pavan Shahapurkar, M. A. Mujtaba, Kiran MC, and Gulam Mohammed Sayeed Ahmed. 2022. “Comprehensive Review on Polymer Composites as Electromagnetic Interference Shielding Materials.” *Polymers and Polymer Composites*. SAGE Publications Ltd. <https://doi.org/10.1177/09673911221102127>.

[15] Wang, Fei, Qianfeng Zhou, Zhe Zhang, Peiwen He, Jiliang Zhang, and Kaiyong Jiang. 2022. “Microwave Absorption Performance of Carbon Black/Poly(lactic Acid) Composite for Fused Filament Fabrication.” *Applied Sciences (Switzerland)* 12 (24). <https://doi.org/10.3390/app122412747>.

[16] Wu, Haihua, Zhenglang Hu, Yutian Li, Peng Qian, Li Liu, and Jianxin Zhou. 2022. “Electromagnetic Absorption Properties and Mechanical Properties of Fe-Ni Alloy/Poly(lactic Acid) Composites Fabricated by Fused Deposition Modeling.” *Fuhe Cailiao Xuebao/Acta Materiae Compositae Sinica* 39 (1): 158–68. <https://doi.org/10.13801/j.cnki.fhclxb.20210311.003>.

[17] Kaseem, Mosab. 2019. “Poly(Lactic Acid) Composites.” *Materials* 12 (21). <https://doi.org/10.3390/ma12213586>.

[18] Bertasius, Povilas, Artyom Plyushch, Jan Macutkevič, Jūras Banys, Algirdas Selskis, Oskars Platnieks, and Sergejs Gaidukovs. 2023. “Multilayered Composites with Carbon Nanotubes for Electromagnetic Shielding Application.” *Polymers* 15 (4). <https://doi.org/10.3390/polym15041053>.

[19] Chen, Xiaojun, Hongyang Liu, Dechao Hu, Huaqing Liu, and Wenshi Ma. 2021. “Recent Advances in Carbon Nanotubes-Based Microwave Absorbing Composites.” *Ceramics International*. Elsevier Ltd. <https://doi.org/10.1016/j.ceramint.2021.05.219>.

[20] Kim, Jin Ah, Dong Gi Seong, Tae Jin Kang, and Jae Ryoung Youn. 2006. “Effects of Surface Modification on Rheological and Mechanical Properties of CNT/Epoxy Composites.” *Carbon* 44 (10): 1898–1905. <https://doi.org/10.1016/j.carbon.2006.02.026>.

[21] Chen, Honghui, Wenle Ma, Zhiyu Huang, Yi Zhang, Yi Huang, and Yongsheng Chen. 2019. “Graphene-Based Materials toward Microwave and Terahertz Absorbing Stealth Technologies.” *Advanced Optical Materials* 7 (8). <https://doi.org/10.1002/adom.201801318>.

[22] Zhang, Hongxia, Bingbing Wang, Ailing Feng, Na Zhang, Zirui Jia, Zhengyong Huang, Xuehua Liu, and Guanglei Wu. 2019. “Mesoporous Carbon Hollow Microspheres with

Tunable Pore Size and Shell Thickness as Efficient Electromagnetic Wave Absorbers.”

Composites Part B: Engineering 167 (June): 690–99.

<https://doi.org/10.1016/j.compositesb.2019.03.055>.

[23] Green, Michael, and Xiaobo Chen. 2019. “Recent Progress of Nanomaterials for Microwave Absorption.” *Journal of Materiomics*. Chinese Ceramic Society. <https://doi.org/10.1016/j.jmat.2019.07.003>.

[24] Liu, Tong, Yu Pang, Mu Zhu, and Satoru Kobayashi. 2014. “Microporous Co@CoO Nanoparticles with Superior Microwave Absorption Properties.” *Nanoscale* 6 (4): 2447–54. <https://doi.org/10.1039/c3nr05238a>.

[25] Ruhiyuddin, M. Z., C. P. Faizul, D. Murizam, and A. R.M. Nazri. 2013. “Study on Milling Periods on the Iron Mill Scale Particle Size and Properties.” In *Advanced Materials Research*, 626:1001–5. <https://doi.org/10.4028/www.scientific.net/AMR.626.1001>.

[26] Singh, Dilip K., P. K. Iyer, and P. K. Giri. 2010. “Diameter Dependence of Interwall Separation and Strain in Multiwalled Carbon Nanotubes Probed by X-Ray Diffraction and Raman Scattering Studies.” *Diamond and Related Materials* 19 (10): 1281–88. <https://doi.org/10.1016/j.diamond.2010.06.003>.

[27] Zhao, Tingkai, Cuilin Hou, Hongyan Zhang, Ruoxing Zhu, Shengfei She, Jungao Wang, Tiehu Li, Zhifu Liu, and Bingqing Wei. 2014. “Electromagnetic Wave Absorbing Properties of Amorphous Carbon Nanotubes.” *Scientific Reports* 4: 1–7. <https://doi.org/10.1038/srep05619>.

[28] Su, Xiao, Ruoyu Wang, Xiaofeng Li, Sherif Araby, Hsu Chiang Kuan, Mohannad Naeem, and Jun Ma. 2022. “A Comparative Study of Polymer Nanocomposites Containing Multi-Walled Carbon Nanotubes and Graphene Nanoplatelets.” *Nano Materials Science* 4 (3): 185–204. <https://doi.org/10.1016/j.nanoms.2021.08.003>.

[29] Gao, Xiaoyan, Jiang Li, Yuan Gao, Shaoyun Guo, Hong Wu, and Rong Chen. 2016. “Microwave Absorbing Properties of Alternating Multilayer Composites Consisting of Poly (Vinyl Chloride) and Multi-Walled Carbon Nanotube Filled Poly (Vinyl Chloride) Layers.” *Composites Science and Technology* 130 (June): 10–19. <https://doi.org/10.1016/j.compscitech.2016.03.004>.

[30] Jorio, A., and R. Saito. 2021. “Raman Spectroscopy for Carbon Nanotube Applications.” *Journal of Applied Physics* 129 (2). <https://doi.org/10.1063/5.0030809>.

[31] Rosdi, Nurshahiera, Raba’ah Syahidah Azis, Ismayadi Ismail, Nurhidayat Mokhtar, Muhammad Misbah Muhammad Zulkimi, and Muhammad Syazwan Mustaffa. 2021. “Structural, Microstructural, Magnetic and Electromagnetic Absorption Properties of Spiraled Multiwalled Carbon Nanotubes/Barium Hexaferrite (MWCNTs/BaFe₁₂O₁₉) Hybrid.” *Scientific Reports* 11 (1). <https://doi.org/10.1038/s41598-021-95332-9>.

[32] Nwigboji, Ifeanyi H., John I. Ejembi, Zhou Wang, Diola Bagayoko, and Guang Lin Zhao. 2015. “Microwave Absorption Properties of Multi-Walled Carbon Nanotube (Outer Diameter 20-30 Nm)-Epoxy Composites from 1 to 26.5 GHz.” *Diamond and Related Materials* 52: 66–71. <https://doi.org/10.1016/j.diamond.2014.12.008>

OPTIMIZATION OF SYNTHESIZING CONDITIONS FOR MXENE (Ti₃C₂) PHOTOCATALYST: EFFECT OF LiF: Ti₃AlC₂ MASS RATIO

Nabilah Saafie¹
Farah Amelia Shahirah Roslan¹
Suriati Sufian^{1,2, *}

¹Chemical Engineering Department, Universiti Teknologi PETRONAS, Seri Iskandar 32610, Malaysia

²Center of Innovative Nanostructures and Nanodevices (COINN), Universiti Teknologi PETRONAS, Seri Iskandar 32610, Malaysia

Abstract: The analysis of the proposed work, specifically the development of the MXene photocatalyst, will be described in this chapter. The results of the analysis are useful for comparing the pristine precursor and etched structures of MXene photocatalyst. A minimally intensive layer delamination (MILD) method will be used to create the MXene photocatalyst which use lower fluorine content solution (HCl-LiF) creates a safer and easier method. The synthesis parameters for the development of highly efficient MXene photocatalysts, such as LiF: Ti₃AlC₂ mass ratio will be optimized. FTIR, XRD, SAP, FESEM and DR UV-Vis analysis will be used to characterize the as-developed MXene photocatalyst and its precursor, Ti₃AlC₂. The hypothesis of this study is etching treatment will increases the hydrophilicity and active functional group such as oxygen (-O), fluorine (F) and hydroxyl (OH) on MXene surfaces. The performance of photocatalytic degradation will be tested with initial dye concentration of 30 ppm, solution pH at pH7 in room temperature ($\pm 27^{\circ}\text{C}$). Each photocatalytic degradation study was performed in 50 ml of methylene blue solution with 0.1 g photocatalyst. All developed MXenes will undergo photocatalytic degradation performance to identify the optimized synthesizing conditions. Highest MXene photodegradation performance was found reached 90% removal within 180 min.

Keywords: MXene, Photocatalytic Degradation, Methylene Blue

1. Introduction

The growth in industrial and agricultural activities in recent years has assigned major responsibility for the continuous discharge of toxic organic contaminants into water bodies without treatment. As a result of the toxic chemicals and gases emitted by industrial and agricultural activities, this growth has been shown to contribute to the rise in water pollution over the century [1]. According to the Natural Resources Defense Council (NRDC), industrial wastewater pollution affects about 1 billion people worldwide each year, with low-income populations being disproportionately affected because their residences are often closer to polluting companies. Furthermore, excessive pollutants such as heavy metals, dyes, and antibiotics were dumped into the water body without sufficient wastewater treatment [2].

Photocatalysis has been shown to be a cheap, easy, efficient, and rapid treatment procedure for complete dye removal using suitable photocatalysts and a light source [3]. In addition, due to the mild conditions and environmentally friendly characteristics, photocatalytic technology has been diffusely investigated by combining the solar energy with photocatalysts in environmental

remediation and energy field, such as the degradation of pollutants in water or air, hydrogen (H₂) production, carbon dioxide (CO₂) reduction and nitrogen (N₂) fixation [4]. Besides, photocatalysis can generate many reactive species that can effectively destroy and degrade organic contaminants [5]. Principally, photocatalysis steps take account of transferring the liquid phase reactants to the catalyst surface, adsorbing the reactants to the catalyst surface, reacting in the absorbing phase, desorbing the final product, and finally removing the final product in the liquid phase.

MXene has captivated researchers interest as a promising environmental photocatalyst, particularly for treating wastewater containing dyes [1]. Mxene materials have gotten a lot of attention because of their wide range of unique structure characteristics, high specific surface area and outstanding physicochemical properties [6]. The MXene compound has a higher reflectivity in the ultraviolet (UV) light correspond to the visible light absorptions with a reported wavelength range of 300 to 500 nm, which is important for photocatalytic reactions with energy gaps of 0.25 to 2.0 eV [7]. MXene also has a hydrophilic character, which is combined with hydrophobicity and an active functional group on its surface, which extracts the adsorptive nature to deal with ionic/molecular species for environmental remediation [8].

Accordance to that, a preliminary synthesis of MXene photocatalyst development were evaluated in this chapter before using it as co-catalyst in heterojunction composite for an optimal manner. The etching procedure resulted in the improvement of hydrophilic compounds due to the functionalization of MXene. However, depending on the precursor mass, etching concentrations, and circumstances, the approaches may reduce photodegradation performance [9]. These also affect the presence of functional groups on MXene surface. Therefore, this present research objectives are to compare the pristine and composite structures of MXene photocatalyst and evaluates their photocatalytic degradation. A minimally intensive layer delamination (MILD) method was chosen which use lower fluorine content solution (HCl-LiF) creates a safer and easier method.

2. Methodology

2.1 Synthesizing the Mxene Photocatalyst

The MXene photocatalyst was developed using the minimally intensive layer delamination (MILD) method [9]. In a typical procedure, 3.2 g LiF was gradually added to 10 ml of 9 M HCl and stirred continuously for 5 minutes. Then, 2.0 g of Ti₃AlC₂ powder was gradually added, and the etching process was allowed to run for 24 hours at room temperature. The as-synthesized resultants were determined as MXene-1.6 which meant the mass ratio of LiF:Ti₃AlC₂. The typical method etching process was repeated for MXene-1.0 and MXene-2.0. The MXene powder was then separated and cleaned with deionized water using a centrifuge process at 3500 rpm for 5 minutes per cycle. The washing procedure continued until the pH of the supernatant was about neutral. The MXene powder particles were then rewashed with 500 ml deionized water by vacuum assisted filtering before being dried at 120°C for 24 hours.

2.2 Characterization of Mxene Photocatalyst

The surface functional groups for MXene photocatalysts were characterized using FTIR spectroscopy (Perkin Elmer Spectrum One, FTIR-Frontier, U.S.A) using KBr pellet method. Their FTIR spectroscopy results were scanned and recorded in the range of 400 to 4000 cm⁻¹. A Powder X-ray Diffractometer was used to analyse the materials using X-ray Diffraction Analysis (XRD) (PANalytical, X'pert Powder, Netherland) in the 2θ angle ranging from 10°

to 80° incremented by 0.02° step size. Field Emission Scanning Electron Microscope (FESEM) (Carl Zeiss, Supra 55 VP Scanning Electron Microscope, U.S.A) was used to examine the structures and surface morphologies of MXene. Based on the nitrogen adsorption-desorption method, the specific surface area, pore volume, and pore size distribution of surface area analyzer and Porosity analysis (SAP) (Micromeritics Tristar II, Surface Area and Porosity, U.S.A) were obtained. The absorption spectrum was estimated using a DRS UV-Vis, UV 3150 NIR (Shimadzu, Japan) spectrophotometer constructed with an integrated sphere.

2.3 Photocatalytic Degradation Analysis

In a photoreactor with a magnetic stirrer, studies on photocatalytic degradation were conducted. An aluminium foil cover will be placed over a 100 mL beaker to prevent radiation leakage. The suspension was positioned 15 cm distant from the beaker to increase visible light irradiation. At pH 7 and 30°C, testing for the photocatalytic degradation of synthetic Methylene Blue (MB) will be conducted. MXene powder (100 mg) was mixed in 50 mL of 30 ppm MB aqueous solution. The suspension was stirred in the dark for 30 minutes to ensure an equilibrium adsorption-desorption state [10]. Following that, cooling fans were used to keep the solution temperature stable, and a 500 W Halogen lamp was used to provide continuous stirring (150 rpm). A sample of 0.1 ml from the solution will be taken using a micropipette every 30 minutes, and it will be examined using a UV-Vis spectrophotometer (UV-1800, Shimadzu).

2.4 Methylene Blue Degradation Analysis

Powdered MB was dissolved in deionized water to create synthetic wastewater. The precise peak of MB's absorption at 664 nm was chosen as the typical peak for concentrations. The following equation will be used to estimate the MB degrading efficiency (%):

$$MB \text{ degradation efficiency} = \frac{C_0 - C_t}{C_0} \times 100 \quad (1)$$

Where C_0 and C_t (mg/L) represents to initial and at certain time concentrations of MB, respectively.

3. Results

3.1 Characterization of Mxene Photocatalyst

As shown in Figure 1, FTIR spectroscopy was utilised to further examine the functional group compositions of Ti3AlC2, MXene-1.0, MXene-1.6, and MXene-2.0. The surface chemistry bands displayed by the FTIR spectra of O–H stretching and bending at 3431 and 1628 cm⁻¹, respectively showed higher transmittance changed for MXenes compared to Ti3AlC2 [11]. This indicates that higher oxygen terminated functional groups were introduced due to the etching process. Since the fluoride salts were used in the etching process, the enhancement of peak intensity of C–F band at 1400 cm⁻¹ was noticed [12]. Presence of new functional group was identified at 571 cm⁻¹ attributed to Ti–O bond vibration, originates from the surface adsorption of –OH [13]. Methylene blue preferentially degrades when an oxygen-containing surface group is present, which is a desired attractive site for cationic dyes [14].

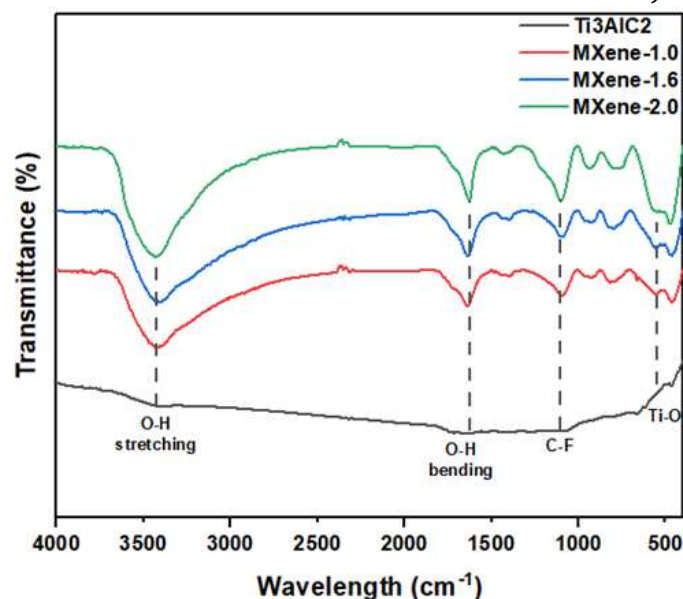


Figure 1: FTIR Spectra

Ti3AlC2, MXene-1.0, MXene-1.6, and MXene-2.0's phase composition was examined using XRD analysis. As shown in Figure 2, all exhibited almost the same XRD patterns, except for some new peaks located at $2\theta = 13.1^\circ$, 45.2° and 65.5° which denoted as Ti3C2 [15]. This proved the transformation of Ti3AlC2 to Ti3C2 causing by LiF etching. The Ti3C2OH's distinctive peaks were seen at $2\theta = 19.3^\circ$, 48.6° , 56.6° , 60.3° and 71.8° which also confirmed the successful fabrication of MXene. However, the intensity peaks of Ti3AlC2 were remained and reduced after treatment due to the incomplete etching using LiF/HCl agents [9]. Table 1 shows the compound composition to assess the score percentage of each phase. After the optimization of synthesizing conditions process MXene-1.6 has the highest Ti3C2 compound score, 69% compared to MXene-1.0 and MXene 2.0.

Table 1: Compound composition score by XRD

Photocatalyst	Percentage (%)		
	Ti3AlC2	Ti3C2	Ti3C2OH
Ti3AlC2	100	NA	NA
MXene-1.0	62	31	7
MXene-1.6	26	69	5
MXene-2.0	31	64	5

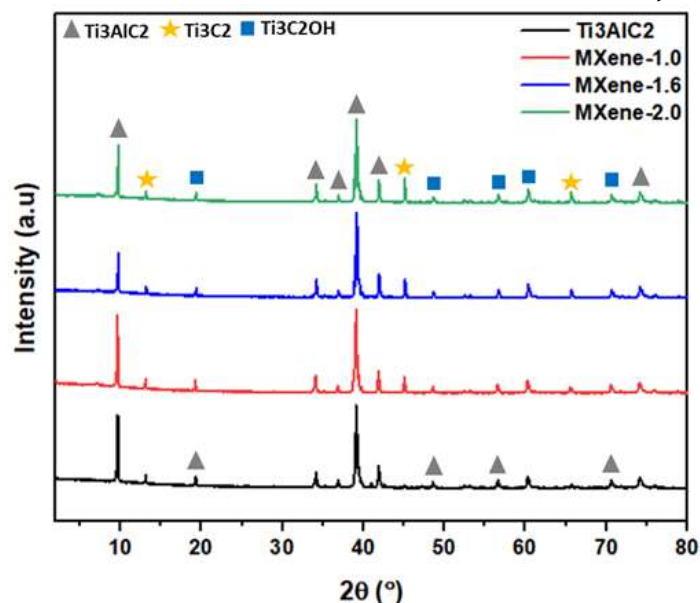


Figure 2: XRD Patterns

The DR-UV-Vis characteristics of the Ti_3AlC_2 , MXene-1.0, MXene-1.6, and MXene-2.0 samples were shown in Figure 3 (a). It is strongly supported that the titanium carbide photocatalyst was a visible-light-driven photocatalyst based on the band gap absorption edge of the samples' resilient absorption capacity in the visible light regime [15]. The band gap energy for the titanium carbide photocatalyst is depicted in Figure 3(b) by extrapolating from the curve of $(\alpha h\nu)^2$ versus $(h\nu)$. Ti_3AlC_2 , MXene-1.0, MXene-1.6, and MXene-2.0 had computed band gaps of 1.37, 1.25, 1.23, and 1.27 eV, respectively. The measured band gap energy agreed with previous findings with band gap energies ranging from 0.25 to 2.0 eV [16].

It is anticipated that photogenerated holes and photoexcited electrons will be easier to move and migrate due to the narrow band gap energy's reduced photon energy need, which is in line with the capacity to absorb visible light. The efficiency of photocatalytic degradation will rise as a result. Additionally, in the heterostructure concept, highly conductive MXenes with known low bandgap energies would make it easier to collect electrons produced by other semiconductors [17]. The migration of generated electrons was favourable to increase separation efficiency of charge carriers and inhibit the electron-hole recombination.

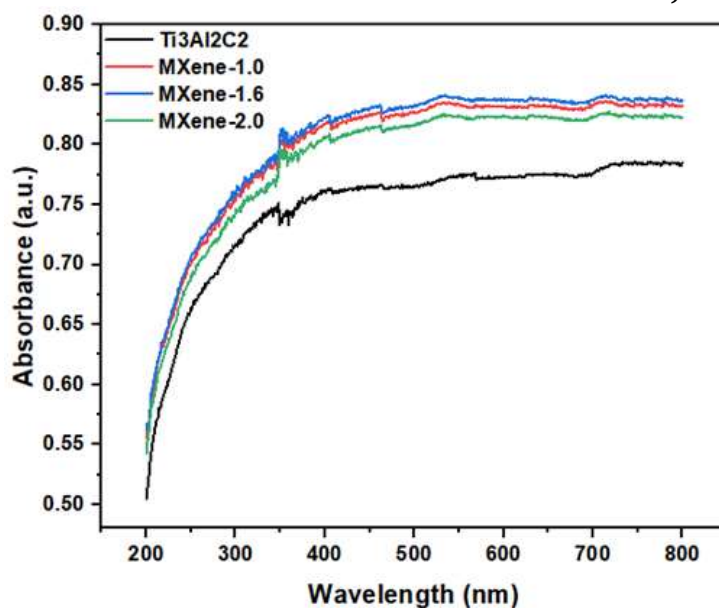


Figure 3: (a) Absorbance Spectra

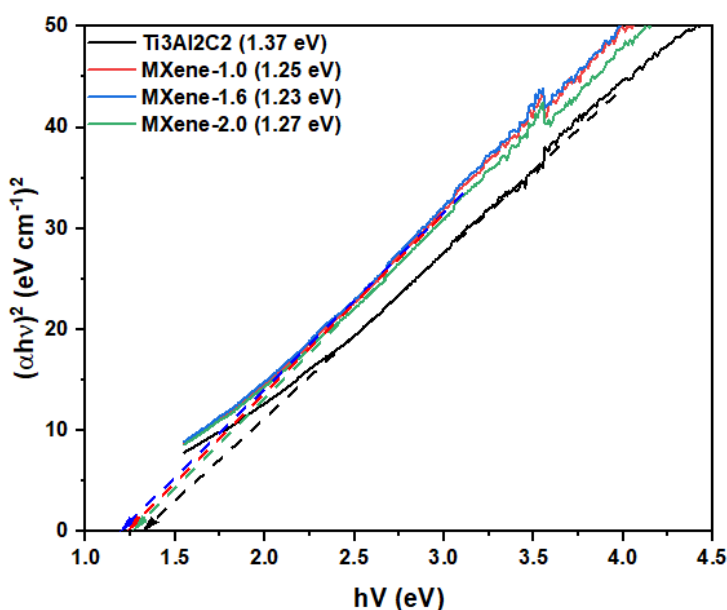


Figure 3: (b) Tauc Plot

The morphological structures of Ti₃AlC₂ and MXene-1.6 are presented in Figure 4 (a, c) and 4 (b, d), respectively. The FESEM images showed that the MXene-1.6 yielded stacked carbon layer with opening lamellas and pores compare to MAX phase, Ti₃AlC₂ appeared as large block morphology. The basic basis for the production of the layered structure, where structural holes can be clearly seen, is the chemical exfoliation of the A group element, Al, during the etching operation [18]. As shown in Table 2, the elemental composition of Al element was plunged from 14.92 wt% in Ti₃AlC₂ to 0.33 wt% in MXene-1.6. This data proved that Al layers were exfoliated, and layer distances expanded. Therefore, Ti₃C₂T_x owing higher surface area which increases the exposed surface functional groups thus expected to improve the photodegradation activity. Furthermore, O element was also increased after etching treatment from 9.78 wt% in Ti₃AlC₂ to 13.01 wt% in MXene-1.6.

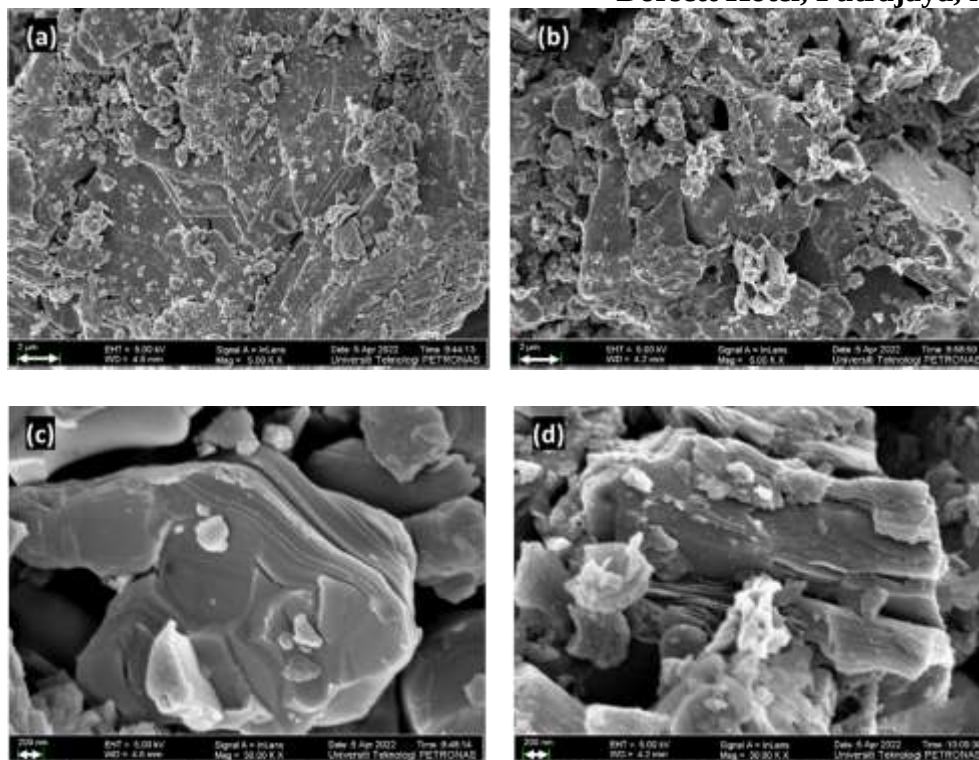


Figure 4: FESEM micrographs of (a, c) Ti₃AlC₂ (b, d) MXene-1.6 at 5.00 K and 30.00 K magnification, respectively

Both Ti₃AlC₂ and MXene-1.6 shows type IV isotherm showing H3 type hysteresis loop in Figure 5 which indicated parallel plate shaped pores [19]. As shown in the image, MXene-1.6 has more volume adsorbed at the isotherm 'knee' than Ti₃AlC₂, indicating that its surface area has increased. This could be explained by the increasing of layer spacing during etching treatment where Al was removed. Table 2 lists the texture metrics, including the Brunauer Emmet Teller (BET) surface area, total pore volume, and average pore diameter.. The result confirmed that the MXene-1.6 had higher surface area 3.11 m²/g compared to the precursor Ti₃AlC₂, 1.68 m²/g. The results obtained agreed with previous work carried out by Qu et al. which reported 2.17 m²/g surface area after HF-forming etchants treatment [13].

Table 2: SAP analysis of the Ti₃AlC₂ and MXene-1.6

Photocatalyst	Surface Area (m ² /g)	Pore Volume (cm ³ /g)	Pore size (nm)
Ti ₃ AlC ₂	1.68	0.01	18.50
MXene-1.6	3.11	0.01	17.10

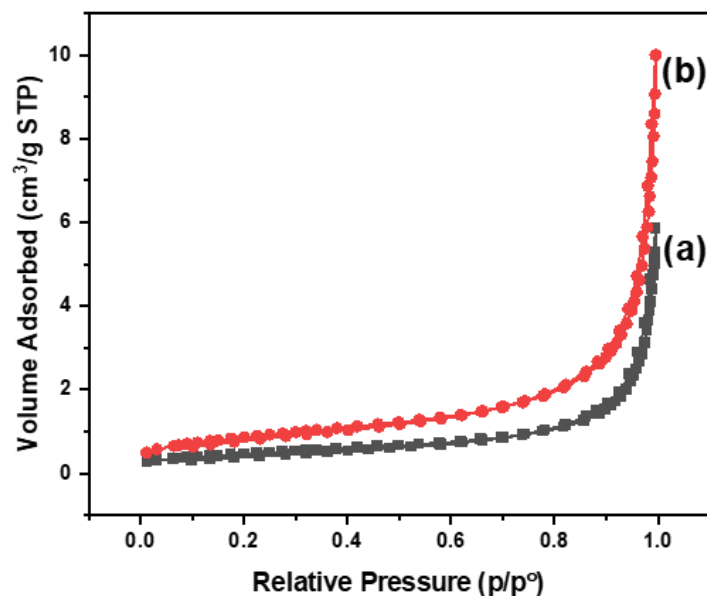


Figure 5: N₂ adsorption-desorption isotherm of (a) Ti₃AlC₂ and (b) MXene-1.6

3.2 B. Photocatalytic Degradation Performance

In a batch setup using MB solutions, the effectiveness of Ti₃AlC₂, MXene-1.0, MXene-1.6, and MXene-2.0 was assessed. The results shows a massive difference between the samples in terms of the amount of time needed to get the dye down to a safe level. The MXene-1.6 shows the highest adsorption rate with 90% of MB being removed within 180 minutes. Meanwhile, MXene-1.0 and MXene-2.0 only managed to removed MB ~45% and ~60%, respectively within the same timeframe. This outcome was agreed with the elemental composition by EDX spot where MXene-1.6 give the higher oxygen terminated functional groups were introduced due to the etching process. MXene 1.6 also has a hydrophilic nature that is combined with hydrophilicity and active functional groups like oxygen (-O), fluorine (F), and hydroxyl (-OH) on its surface, which extracts the adsorptive nature to deal with ionic/molecular species for environmental resolution [19].

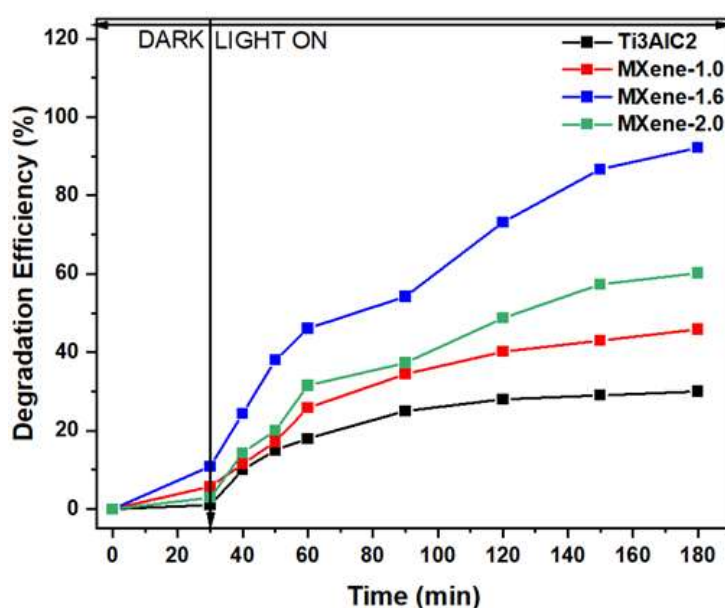


Figure 6: Decolorization of MB (30 ppm)

4. Conclusions

In this work, MXene-1.6 showed the best LiF:Ti₃AlC₂ mass ratio to synthesize MXene photocatalyst. It offers an adequate number of significant functional groups to connect to the surface of MXene during the etching process. The XRD data revealed that the transformation of Ti₃AlC₂ to Ti₃C₂ causing by LiF etching was successful and MXene-1.6 has the highest Ti₃C₂ compound score, 69% compared to MXene-1.0 and MXene 2.0. The greater BET surface area of MXene 1.6 (3.11 m²/g) compared to the precursor Ti₃AlC₂ (1.68 m²/g) was further validated by a FESEM micrograph. The adsorption analysis indicates that MXene-1.6 has the maximum degradation rate, removing 90% of MB in about 180 minutes. Therefore, a minimally intensive layer delamination (MILD) method successfully created the MXene photocatalyst using lower fluorine content solution (HCl-LiF) with 1.6 of LiF:Ti₃AlC₂ mass ratio.

Acknowledgement

This research was funded by a grant from Ministry of Higher Education of Malaysia (FRGS/1/2020/TK0/UTP/02/22) and Yayasan Universiti Teknologi PETRONAS, YUTP (015LC0-357).

References

- [1] B. M. Jun, S. Kim, H. Rho, C. M. Park, and Y. Yoon, "Ultrasound-assisted Ti₃C₂Tx MXene adsorption of dyes: Removal performance and mechanism analyses via dynamic light scattering," *Chemosphere*, vol. 254, Sep. 2020, doi: 10.1016/j.chemosphere.2020.126827.
- [2] M. Rethinasabapathy et al., "Iron oxide (Fe₃O₄)-laden titanium carbide (Ti₃C₂Tx) MXene stacks for the efficient sequestration of cationic dyes from aqueous solution," *Chemosphere*, vol. 286, Jan. 2022, doi: 10.1016/j.chemosphere.2021.131679.
- [3] M. Yu, H. Liang, R. Zhan, L. Xu, and J. Niu, "Sm-doped g-C₃N₄/Ti₃C₂ MXene heterojunction for visible-light photocatalytic degradation of ciprofloxacin," *Chinese Chemical Letters*, vol. 32, no. 7, pp. 2155–2158, Jul. 2021, doi: 10.1016/j.ccllet.2020.11.069.
- [4] W. Liu, M. Sun, Z. Ding, B. Gao, and W. Ding, "Ti₃C₂ MXene embellished g-C₃N₄ nanosheets for improving photocatalytic redox capacity," *J Alloys Compd*, vol. 877, Oct. 2021, doi: 10.1016/j.jallcom.2021.160223.
- [5] Y. Zhou, M. Yu, R. Zhan, X. Wang, G. Peng, and J. Niu, "Ti₃C₂ MXene-induced interface electron separation in g-C₃N₄/Ti₃C₂ MXene/MoSe₂ Z-scheme heterojunction for enhancing visible light-irradiated enoxacin degradation," *Sep Purif Technol*, vol. 275, Nov. 2021, doi: 10.1016/j.seppur.2021.119194.
- [6] Y. Cui et al., "Biomimetic anchoring of Fe₃O₄ onto Ti₃C₂MXene for highly efficient removal of organic dyes by Fenton reaction," *J Environ Chem Eng*, vol. 8, no. 5, Oct. 2020, doi: 10.1016/j.jece.2020.104369.
- [7] K. Kannan, K. K. Sadasivuni, A. M. Abdullah, and B. Kumar, "Current trends in MXene-based nanomaterials for energy storage and conversion system: A mini review," *Catalysts*, vol. 10, no. 5. MDPI, May 01, 2020. doi: 10.3390/catal10050495.
- [8] K. Huang, C. Li, L. Wang, W. Wang, and X. Meng, "Layered Ti₃C₂ MXene and silver co-modified g-C₃N₄ with enhanced visible light-driven photocatalytic activity," *Chemical Engineering Journal*, vol. 425, Dec. 2021, doi: 10.1016/j.cej.2021.131493. [7] M. Alhabeb et al., "Guidelines for Synthesis and Processing of Two-Dimensional

- Titanium Carbide (Ti₃C₂T_x MXene)," Chem. Mater., vol. 29, no. 18, pp. 7633–7644, 2017, doi: 10.1021/acs.chemmater.7b02847.
- [9] M. Alhabeb et al., "Guidelines for Synthesis and Processing of Two-Dimensional Titanium Carbide (Ti₃C₂T_x MXene)," Chem. Mater., vol. 29, no. 18, pp. 7633–7644, 2017, doi: 10.1021/acs.chemmater.7b02847.
- [10] N. Liu, N. Lu, Y. Su, P. Wang, and X. Quan, "Fabrication of g-C₃N₄/Ti₃C₂ composite and its visible-light photocatalytic capability for ciprofloxacin degradation," Sep. Purif. Technol., vol. 211, pp. 782–789, Mar. 2019, doi: 10.1016/j.seppur.2018.10.027.
- [11] C. Peng, X. Yang, Y. Li, H. Yu, H. Wang, and F. Peng, "Hybrids of Two-Dimensional Ti₃C₂ and TiO₂ Exposing {001} Facets toward Enhanced Photocatalytic Activity," ACS Appl. Mater. Interfaces, vol. 8, no. 9, pp. 6051–6060, Mar. 2016, doi: 10.1021/acsami.5b11973.
- [12] I. A. Alsafari, "Synthesis of CuO/MXene nanocomposite to study its photocatalytic and antibacterial properties," Ceram. Int., no. December 2021, pp. 1–9, 2022, doi: 10.1016/j.ceramint.2021.12.315.
- [13] J. Qu, D. Teng, X. Zhang, Q. Yang, P. Li, and Y. Cao, "Preparation and regulation of two-dimensional Ti₃C₂T_x MXene for enhanced adsorption–photocatalytic degradation of organic dyes in wastewater," Ceram. Int., vol. 48, no. 10, pp. 14451–14459, May 2022, doi: 10.1016/j.ceramint.2022.01.338.
- [14] N. Saafie, M. F. R. Samsudin, S. Sufian, and R. M. Ramli, "Enhancement of the activated carbon over methylene blue removal efficiency via alkali- acid treatment Enhancement of the Activated Carbon over Methylene Blue Removal Efficiency via Alkali-Acid Treatment," vol. 020046, no. July, 2019.
- [15] W. Liu, M. Sun, Z. Ding, B. Gao, and W. Ding, "Ti₃C₂ MXene embellished g-C₃N₄ nanosheets for improving photocatalytic redox capacity," J. Alloys Compd., vol. 877, p. 160223, 2021, doi: 10.1016/j.jallcom.2021.160223.
- [16] J. K. Im et al., "Review of MXene-based nanocomposites for photocatalysis," Chemosphere, vol. 270, May 2021, doi: 10.1016/j.chemosphere.2020.129478.
- [17] X. Sun et al., "Few-layer Ti₃C₂T_x (T = O, OH, or F) saturable absorber for a femtosecond bulk laser," Opt. Lett., vol. 43, no. 16, p. 3862, Aug. 2018, doi: 10.1364/ol.43.003862.
- [18] Z. H. Jaffari, S. M. A. Abuabdou, D. Q. Ng, and M. J. K. Bashir, "Insight into two-dimensional MXenes for environmental applications: Recent progress, challenges, and prospects," FlatChem, vol. 28, no. May, p. 100256, 2021, doi: 10.1016/j.flatc.2021.100256.
- [19] R. Long et al., "Ti₃C₂ MXene/NH₂-MIL-88B(Fe): Research on the adsorption kinetics and photocatalytic performance of an efficient integrated photocatalytic adsorbent," Appl Surf Sci, vol. 570, Dec. 2021, doi: 10.1016/j.apsusc.2021.151244.

A FRAMEWORK STUDY ON A REAL-TIME OPERATED SYSTEM FOR EMERGENCY MEDICAL SERVICES (EMS) USING NODEMCU PROCESSOR

Azwati Azmin^{1*}

Samihah Abdullah¹

Zafirah Faiza¹

Najwa Rawaida Ahmad @ Ahmad Fauzi²

Rabiatuladawiah Akhbar²

Wan Mohd Yusof Rahiman³

¹ School of Electrical Engineering, College of Engineering, Universiti Teknologi MARA, Cawangan Pulau Pinang, Malaysia

² Centre of Foundation Studies, Universiti Teknologi MARA, Cawangan Selangor, Kampus Dengkil, Dengkil 43800, Selangor, Malaysia

³ School of Electrical and Electronic Engineering, Universiti Sains Malaysia

Abstract: *Emergency Medical Services (EMS) has always been a great help in treating patients before they arrive at the hospital. On the route to the hospital, where they will receive proper and additional treatment, pre-hospital care is an emergency aid to help stabilise the patient's condition. However, there is a lack of effective maintenance in the data interchange between the emergency department doctors and the paramedics in the ambulance. To give the patient the highest level of treatment, the patient's handoffs from paramedics or pre-hospital to in-hospital staff are not of sufficient quality. Hence came the idea of incorporating the Internet of Things into the emergency healthcare system. This paper discusses a proposed idea to design a reliable healthcare monitoring system that enables data exchange between doctors and paramedics efficiently. This system utilises the IoT platform, Blynk App, to display the data in both places (ambulance and emergency department) by using NodeMCU as the core processor and wearable sensors to measure the main parameters and vital sign data (as is current practise in ambulances). This system may also confirm that data from the sensors is transmitted in real-time to the host computer. The findings indicate that the vital sign readings are roughly accurate when compared to actual devices or marketable measurement tools. At the same time, the system will help the paramedics and doctors communicate efficiently in two ways and enable them to exchange patient details in real-time, so appropriate action can be discussed and the patient can be treated swiftly. The provided information allows hospital professionals in the emergency department to suggest what tools or medications should be ready to treat the patient further.*

Keywords: *Emergency Medical Services, Healthcare Monitoring, Mobility Solutions, Internet of Things, NodeMCU*

1. Introduction

Technology has been implemented in almost every aspect in our daily life to an extent that it plays an important role in today's world. Technology is the dynamic application of techniques, methods, skills, and processes utilized for practical purposes in the creation of goods and services or in accomplishing outcomes that benefits consumers [1]. This also applies to

healthcare industries where patients need for the absolute care is a precedence. One of the many opportunities exist to improve the quality of a service or industries is to increase the use of information technology in the said field. Although the combination of Internet of Things (IoT) and medical is not uncommon, it has a lot of potential to improve and incorporate better technology to assist both the emergency healthcare and the healthcare department.

Incorporating technology in emergency healthcare would tremendously help to further enhance the quality of patient's care. This focuses more on the patients that are being rushed to hospitals from incident site in an ambulance as the urgency to treat their wounds falls into the paramedic's priority when it comes to keeping them alive. The condition of the patient may deteriorate along the way to the hospital if they are not assessed and treated in time. Application of technology into this field could reduce the number of fatalities occurred. Countless amount of research has been done to search for advanced, accurate and systematic technology for a better healthcare industry.

The world has been shaken with a new deadly virus that causes a public health emergency to be declared by WHO on 30th January [2]. The virus namely Coronavirus disease 19 or better known as COVID-19 has spread all over the world raising public panic and chaos. The outbreak was denoted as a pandemic seeing it has already spread across the entire world in months and it continues to haunt us to this day. The hospitals and the hospital staffs had it worse as many people were being admitted to hospitals daily requiring emergency care. The rising cases of covid-19 which had caused quite a traumatic impact to the world especially the hospital staffs as they need to treat the patient coming in and out of the hospital daily without rest. With the inpatient that have been in the hospital even before covid-19, the number of patients has increased immensely to the point that the hospital started to run out of hospital beds. The hospital staffs were pushed to make decision quickly to save more patients.

The application of Internet of Things (IoT) into emergency healthcare could be the better digital solutions to saving lives in the time of rush. The patient's vital signs, TRIAGE tags [3] and conditions could be updated to the doctors in the emergency department without lag. Both parties can also communicate with one another to discuss the patient's treatment. This welcomes more possibilities in telemedicine and disaster management. The main purpose of this research is to find the optimum way to improve patient's care once arrived in emergency department. The rest of this paper is organized as follows: Section II describes the work done related to using different processor to collect vital signs data in efficient way. It also gives a comparison using different processor versus data transfer for public health monitoring. Section III describes the methodology of proposed framework implementation of wearable health monitoring system mimics from the current devices are use in ambulance. The replacement wearable sensors are limited to measure the vital sign data connected to the NodeMCU processor and built-in with Wi-Fi module inside to establish the communication between processor and IoT platform to display the data. Section IV describe all the findings from the circuit connections and test phase of this system and discussion was made when comparing the actual data taken with the approved equipment used to determine the hardware accuracy. Section V discusses the objectives of this system based on the performance evaluation results and some points of future recommendations to improve the credibility of health system monitoring.

2. Literature Review on Related Work

Several research on past papers related to health monitoring have been conducted through explorative method of review that led to the discovery of technology in emergency healthcare by reviewing various related articles. Some of which provided a whole new perspective and insights on the topic itself. The review aims to distinguish the various viewpoint that exist on the implementation of mobility solutions in emergency healthcare. Ashikur Rahaman [4] have gathered a few papers as reviews that discusses on the IoT based health monitoring system. The systems gathered ranges in the microprocessor that they employed, the uses, cost, and the feedback device of the system. It can be concluded that although IoT is considered to be hard to implement in a system, it was undoubted that IoT in the medical world would bring great benefit to both the patient and doctors as they helped in informing real time data. IoT also allows the monitoring process to be done remotely. Early detection of health problems allows the doctors to take immediate action to potentially help save the patient's life.

According to Jayakumar [5], the concept of IoT helps in monitoring the patient's health efficiently. He suggested a system that uses Arduino UNO as its microprocessor, heart rate sensor, ECG monitor sensor, and LM35 temperature sensor to be used as the sensors that will monitor the patient's health. Their project was designed to be minimalistic and effective by sensing and transferring data to the webpage provided for every 20 seconds. Once the patient's parameters recorded a value that exceeds the threshold value, doctors will be alerted via a message. The system uses a NodeMCU to help transfer the data acquired to a website that can be displayed from either a PC or a mobile. It employs the software Thingspeak to receive and display the data for monitoring purposes. Mehmet Tastan [6], discusses of a wearable smart health monitoring system incorporating an android based application. The system incorporates Arduino Pro mini as its microcontroller with a few sensors such as pulse sensor, and a temperature sensor to gather the patient's heartbeat and temperature to monitor. The system uses a Bluetooth module to transfer the data to blynk that will display the heartbeat and temperature. When an emergency occurred typically whenever the parameters recorded went over the threshold value, a notification will be sent to the doctor and family members. This helps to rescue the patient in time since the data is being displayed in real time.

Ranjeet Kumar [7] has proposed a system of an IoT based health monitoring system using android app. The system incorporates a BSN which refers to a body sensor network and other smart products to achieve its aim which was to improve the efficiency and quality of healthcare. The sensors used was pulse sensor and DHT11 that will measure the patient's heartbeat and temperature together with humidity respectively. The system installed a raspberry pi as its microcontroller that will cater to the commands and run the system as desired. The system has an application that will receive the data transferred to be then displayed of the values of the patient's temperature and heartbeat. The system was set so that whenever the values for the temperature sensor went above 38 degrees Celsius or below than 37 degrees Celsius. The family members and the doctors will be informed depending on the seriousness of the situation. Prof. M.P. Sardey [8] suggested an IoT based patient health monitoring system that allows for continuous monitoring of the patient's vital sign. The proposed system has a LPC2148 as its microcontroller, temperature sensor and heart rate sensor as the input that will measure the patient's measurement. In addition, it uses a liquid crystal display with a measurement of (16x2) to display the measurements taken of the patient. It employs a GPRS which stands for General Packet Radio Service, Protocol that enables users to access the data at any time of the day. The system proposed also uses MAX232 which is a dual driver that converts TTL level to RS232 level. A buzzer was installed to help alert the people nearby of any dangerous condition.

Another study by Rathy G. A. [9], who proposed an idea of an efficient IoT based biomedical health monitoring and diagnosing system using myRIO to add twist to the idea of emerging IoT with medical world. The proposed system employs heart rate sensor, pulse sensor, BP sensor wearable ultrasound patch, MAX30205 temperature sensor as the sensors that will measure the patient's measurement, ESP 8266 as a Wi-Fi transceiver development kit, and NI-myRIO as a hardware to process real-time data. The system uses LabVIEW to display the readings taken and monitor the patient's health. Once the sensors measure the readings from the patient, it will transfer the data through Wi-Fi module and display them on the LabVIEW. It also equips an alert system that will alert the caretaker and patient to either take their medicine or to check on the patient.

Emergency medical services which is pivotal to the healthcare system is an aid given to inpatient and outpatient of hospitals which is necessary to prevent death or serious impairment of the health of the patient. During emergency treatment, the paramedics will first assess the patient's condition before proceeding with appropriate treatment. Emergency assessment is a process that is used to classify and treat life-threatening problems. It assessed the level of consciousness, cervical spinal stabilization, airway, breathing, and circulation [10]. The paramedics are always on the ready to give medical services or emergency medical procedure for the patient that they are attending to in the ambulance. Furthermore, it is necessary to keep the patient alive during the transportation of the patient from incident scene to the hospital [11].

2.1 Development of Framework Study for Emergency Medical Services

Emergency medical services also is an extensive system which supplies the disposition of personnel, facilities, and equipment to ensure efficient, systematic, and prompt delivery of health and safety services to patients of sudden injury, illness, and trauma thus, the incorporation of emergency medical services is fervently endorsed by the World Health Organization (WHO) as one of the crucial components of the healthcare system [12]. It can also help in reducing the rates of suicide [13]. Emergency assessment is an essential step in taking care of the patient hence improving the assessment of the emergency healthcare would reduce fatalities when transporting the patient in an ambulance.

Ambulance is considered as part of the emergency medical services in a form of transportation which is provided for the reason of taking care of sick and injured people to and from hospital usually in the case of emergency and paramedics will tend to the patient in the ambulance. Paramedics, emergency medical technician, emergency medical personnel, emergency dispatch personnel and call-takers are a part of the emergency response team, and their role is to coordinate the response communication for pre-hospital and out of hospital medical emergency care [14]. Not only do the paramedics in the ambulance need to have the medical knowledge on emergency care, but the ambulance also requires a reliable communication system with a back-up system to enable a smooth communication [15]. The paramedics must know how to calm the patient and talk them through the treatment all the way to the hospital to keep the patient relaxed while they treat them therefore, by allowing the paramedics to communicate with the doctors will ensure the paramedics of the treatment executed on the patient [16]. It was shown that communication is crucial in be it between the paramedics and the hospital staffs or with the patient to guarantee that the patient is alive, and their health and safety are taken care of. Enhancing the quality of communication system between the paramedics and the hospital staffs would greatly help in reducing casualties caused by delayed communication. Health information exchange is a patient's summary which consists of

previous history, physical examination, diagnosis, treatments, and laboratory investigation [17].

2.2 Parameters Taken During Emergency Medical Service

Vital signs consist of systolic blood pressure, pulse pressure, heart rate, respiratory rate, and Glasgow coma scale [18]. Vital signs are the key to help treat the patient in early identification of distress, wounds or diseases that poses threat to the patient's health causing it to deteriorate along the way that could later be fatal [19]. Vital signs are closely related to TRIAGE and helps in determining the TRIAGE tags for each patient. Important vital signs are blood pressure, SpO₂, heart rate, respiration rate and temperature of the patient. Blood pressure is divided into two categories which are systolic and diastolic. Systolic blood pressure shows the pressure exerted on the artery walls when the heart beats and diastolic blood pressure defines the pressure deployed to the artery walls when the heart is in resting condition which is between the heart beats. Normal values for blood pressure should be less than 120/80 mm/Hg. The former value holds for the systolic blood pressure value meanwhile the latter value denoted as the diastolic blood pressure value. If the blood pressure is higher than the normal value, it indicates hypertension.

Next, SpO₂ is oxygen saturation which is the measurement of the amount of oxygen carrying hemoglobin in the blood in relative to the amount of hemoglobin that is not oxygen carrying. The oxygen in the blood needs to be at certain level for the body to function properly. The normal SpO₂ level should be in the range of 95% to 99%. Furthermore, heart rate is also known as pulse. It is the number of heart beats per minute. The heart rate may vary each person though the heart rate tends to be lower when one is resting, and it can increase when the person is moving or exercising. The normal range for the heart rate is quite different between an adult and a child. The normal range for a child range from 70 beats to 100 beats per minute while adult's heart rate ranges from 60 beats to 100 beats per minute.

There is a way to count the heart rate which is to patiently take the pulse on the wrist or neck for about 10 seconds then multiple it by 6. Maximum heart rate can also be calculated via simple mathematical solution. The predicted maximum heart rate can be calculated by deducting 220 from your age. Moreover, respiration rate defines as the number of breaths that the person takes per minute. Respiration rate is taken when the patient is in resting condition. It may increase due to illness, fever, and other medical conditions. The normal respiration rate for an adult is within the range of 12 breaths to 16 breaths per minute. Lastly, temperature is one of the most pivotal signs of the things happening inside the person. Body temperature may vary from person to person and types of activities done throughout the day affects the body temperature. Body temperature could be taken by the mouth, armpit, forehead, ear, and rectum. Normal body temperature for a human is 37 degrees Celsius. All the possibilities readings for vital signs are shown in Fig. 1 – 5 below [16-19].

BLOOD PRESSURE CATEGORY	SYSTOLIC mm Hg (upper number)	and/or	DIASTOLIC mm Hg (lower number)
NORMAL	LESS THAN 120	and	LESS THAN 80
ELEVATED	120 - 129	and	LESS THAN 80
HIGH BLOOD PRESSURE (HYPERTENSION) STAGE 1	130 - 139	or	80 - 89
HIGH BLOOD PRESSURE (HYPERTENSION) STAGE 2	140 OR HIGHER	or	90 OR HIGHER
HYPERTENSIVE CRISIS (consult your doctor immediately)	HIGHER THAN 180	and/or	HIGHER THAN 120

Fig. 1. Blood Pressure Readings and Indications.

Blood oxygen saturation (SpO ₂)			
100 - 98 %		Normal	
97 - 95 %		Insufficient Tolerable, patient hardly notices any influence	
94 - 90 %		Decreased Immediate intervention (eating, exercise)	
< 90 %		Critical Referral to specialist	
< 88 %		Severe hypoxia Hospitalization	
< 78 %		Acute danger to life	

Fig. 2. SPO2 Readings and Indications.

Age (years)	Respiratory rate (per minute)	Heart rate (per minute)
<1	30-40	110-160
1-2	25-35	100-150
2-5	25-30	95-140
5-12	20-25	80-120
>12	15-20	60-100

Fig. 3. Heart Rate and Respiration Rate Readings Divided by Age.

	Measurement site		
	Mouth / armpit	Ear / forehead	Rectum
Low temperature	< 35.8	< 35.7	< 36.2
Normal temperature	35.9 - 37.0	35.8 - 36.9	36.3 - 37.5
Increased temperature	37.1 - 37.5	37.0 - 37.5	37.6 - 38.0
Light fever	37.6 - 38.0	37.6 - 38.0	38.1 - 38.5
Moderate fever	38.1 - 38.5	38.1 - 38.5	38.6 - 39.0
High fever	38.6 - 39.5	38.6 - 39.4	39.1 - 39.9
Very high fever	39.6 - 42.0	39.5 - 42.0	40.0 - 42.5

Fig. 4. Temperature Readings on Site of The Temperature Taken.

Low temperature	Consult a doctor.
Normal temperature	You are perfectly well.
Increased temperature	You should get some rest.
Light fever	Check your temperature regularly and rest.
Moderate fever	Check your temperature regularly. Consult a doctor if you get worse or if the fever lasts for more than three days.
High fever	Consult a doctor, especially if the fever lasts for more than one day.
Very high fever	Go to emergency ward of a hospital!

Fig. 5. Indications for The Temperature Readings.

2.3 Process Flow of the Vital Sign Monitoring System

The communication in emergency medical services is utterly important because of the high-risk nature of patient handling and patient handoffs to medical error. In cases where the patients were being brought to the emergency department by an ambulance, the only person with the most medical information on the patient in that time is the paramedics. Paramedic's records critical information while tending to the patient in the ambulance to elevate their chance of survival. This information is vital to the hospital staffs that is going to handle the patient afterwards and this could greatly affect the healthcare and triage options. When an accident occurs, the emergency medical service response starts by receiving a report of the incident which will be relayed with the use of an in-station alerting system, pagers, or other means of communication to notify the emergency medical service teams of a need for their service.

Then, the team will acknowledge the receipt of the call by repeating the information received back to the dispatcher to ensure the responding team have the correct information. The responding team should always notify the dispatcher when in route to the location of the incident and all the events happening whilst on the road to enable the dispatcher to dispatch another team if the existing team runs into a problem that may delay the journey. In the ambulance is where the first aid treatment of the patient and health information exchange occurs. Health information exchange is a patient's summary which consists of previous history, physical examination, diagnosis, treatments, and laboratory investigation.

Emergency medical services utilizes the radio to exchange health information, patient's details, and the condition of the patient with the hospital staffs in the emergency department. Sometimes the information is not available or is received at such little data or even incorrect and this causes the diagnosis and treatment of the patient to be prolonged when arrived at the emergency department. The radio communication system used may experience interference with other electrical devices and surrounding environment. Internet of Medical Things gave access to a continuous, remote, and real-time monitoring of patients. Although there are plenty of obstacles to overcome relating to communication protocols, flexibility, bandwidth, reliability, big data and data volume, data management, energy efficiency as well as data security and, privacy [20].

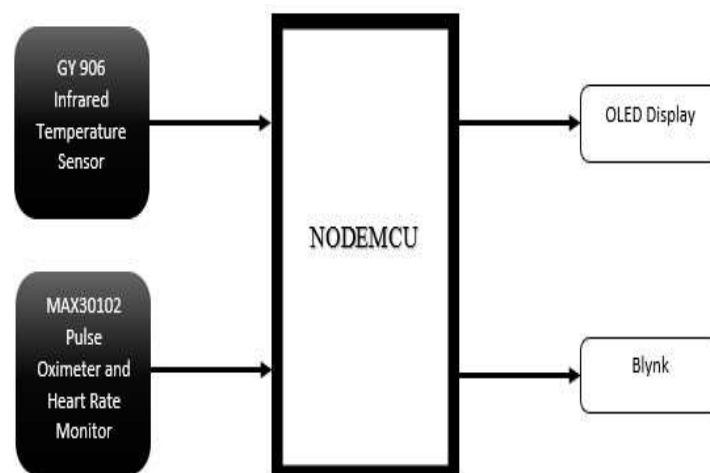


Fig. 7. Block diagram of the Process Flow.

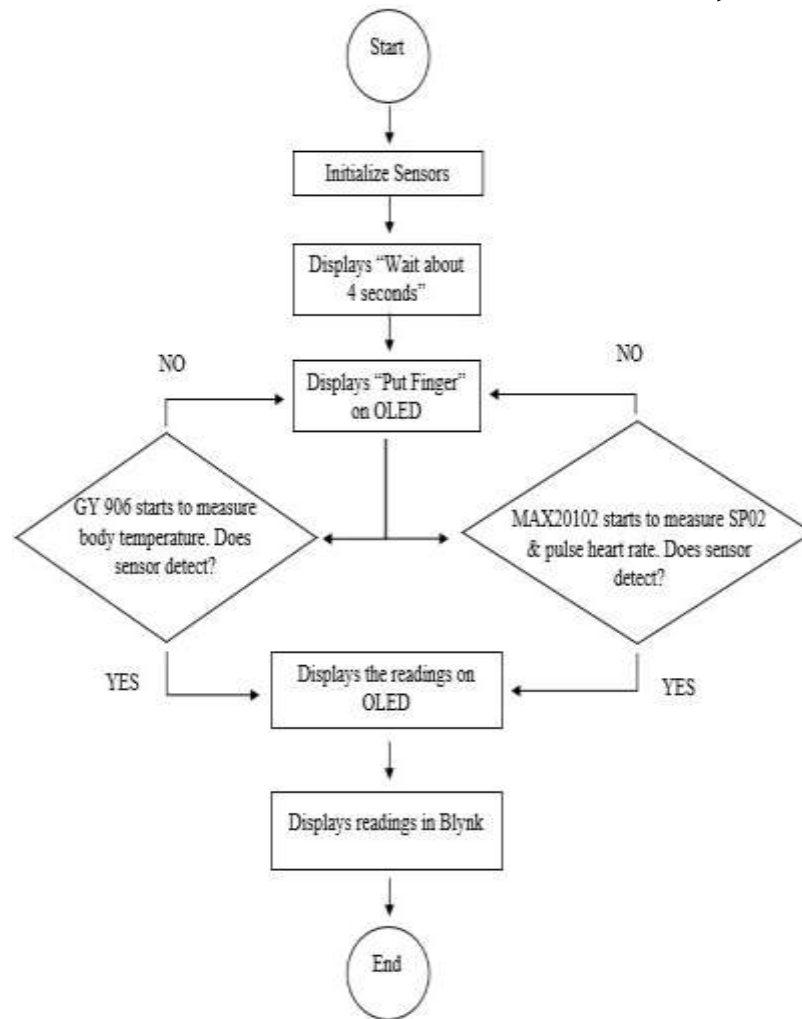


Fig. 8. Flowchart of the Project.

The block diagram of patient's vital sign monitor used in Emergency Medical Services specifically ambulance shown in Fig. 7 above. The block diagram consists of two inputs, one microcontroller and two outputs. Since the microcontroller chosen to be used in the project is NodeMCU, Wi-Fi module can be omitted from the project. MAX30102 and GY 906 is considered as input of the system which will be controlled by NodeMCU that acts as the system's microcontroller. NodeMCU will then transfer the data to the output of the system which are Serial Monitor, OLED, and Blynk. This system will operate whenever the sensors are attached to a patient since it will start to measure the readings. The data read will then be displayed on a screen in the ambulance for the paramedics to monitor the patient and simultaneously display the same data in real-time on a screen in the emergency department of a hospital for the doctors and nurses to monitor. The vital signs gathered from the systems are heart rate, SPO2 level indicating the oxygen level in the blood and the body temperature of the patient. The overall flowchart of this project is shown by Fig. 8. The aim of this project is to transfer the data from ambulance monitor to the monitor in the emergency department.

3. Results and Discussion

The process starts when the system is turned on. The sensors will then be initialized and ready to take measurements from a patient. Once there is a presence on the sensors, it will start to measure the readings. In an ambulance, once the paramedics attach the sensors to the patient, it will start to measure the patient's body temperature, SPO2 level, and pulse heart rate. The readings will then be displayed on OLED or a screen in the ambulance for the purpose of monitoring the patient on behalf of the paramedics. Simultaneously, the same real-time readings will be displayed on blynk or a screen in the emergency department. The doctors and nurses in the emergency department will be able to monitor the patient's vital signs from afar allowing them to prepare for equipment's needed, rooms, apparatus needed to welcome the patient during handoffs. Once the ambulance arrive at the emergency department, the doctors and nurses will have to continue to monitor the patient but all the hassle of preparing the room will be reduced as it had been done before the patient arrive saving time and helps in increasing the patient's chance in surviving and improving the patient's care in the hospital. Fig. 9 shows the parameters shown on OLED when the system was started.



Fig. 9. The Parameters Shown on OLED.



Fig. 10. The Measurements Displayed in Serial Monitor.



Fig. 11. The Parameters Shown on Blynk Appearance.

The serial monitor that displays the measurements of the heart rate, the SpO2 and the body temperature was shown in Fig. 10. The serial monitor can be a screen to display the measurements, and this was used to compensate for the small OLED screen which only showed heart rate and SpO2 values. The system delayed for about four seconds to synchronize the data collected from the sensors to be displayed on all three screens. The measurements collected from the MAX30102 pulse oximeter and heart rate biosensor is a bit fast for the OLED to display the measurements, hence the delay was needed. The values of the heart rate can also be observed to have been increased a little over the normal range which means that the sensors also detect the abnormal values for the measurements making it reliable to be used for human's vital sign measurements. It was also observed that the system can take measurements and display the data continuously for a long period of time.

Table1: The Data Measured from Approved Equipment and From Hardware of The Project.

Condition	Measurements					
	Average recorded parameters			Hardware recorded parameters		
	Heartbeat (bpm)	SpO2 (%)	Body Temperature (°C)	Heartbeat (bpm)	SpO2 (%)	Body Temperature (°C)
Resting	90	99	36.8	90	99	36.8
(children)	83	99	37.3	83	99	37.4
(adult)	94	97	37.8	94	96	37.8
Sleeping	60	100	36.4	61	99	36.5
Jogging	115	97	37.2	115	97	37.2
Running	152	95	37.4	152	96	37.4
Dancing	119	97	37.1	118	97	37.1
Exercising	131	96	36.9	131	96	36.8
Playing basketball	120	96	37.5	120	96	37.5

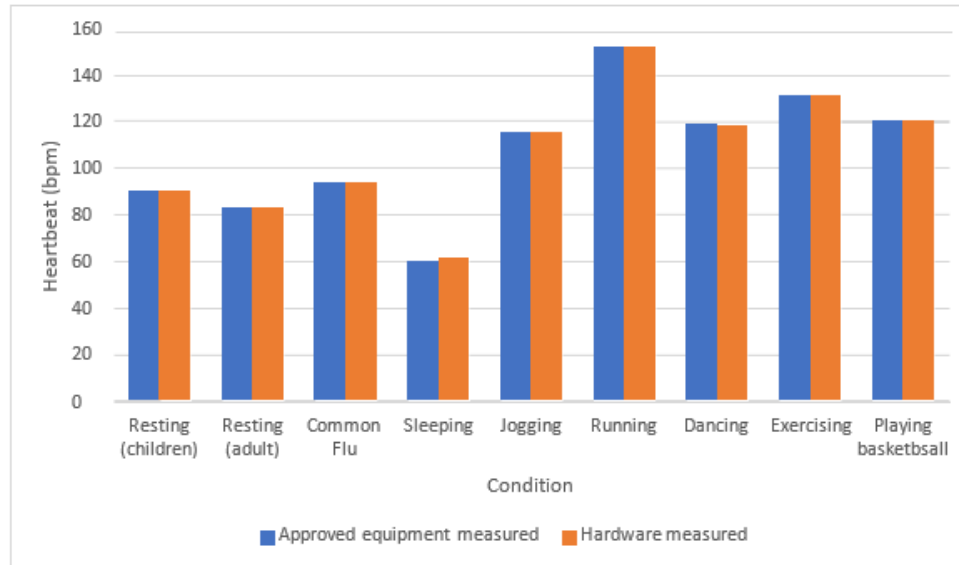


Fig. 12. The Heartbeat Per Minute Taken by Approved Equipment and Hardware.

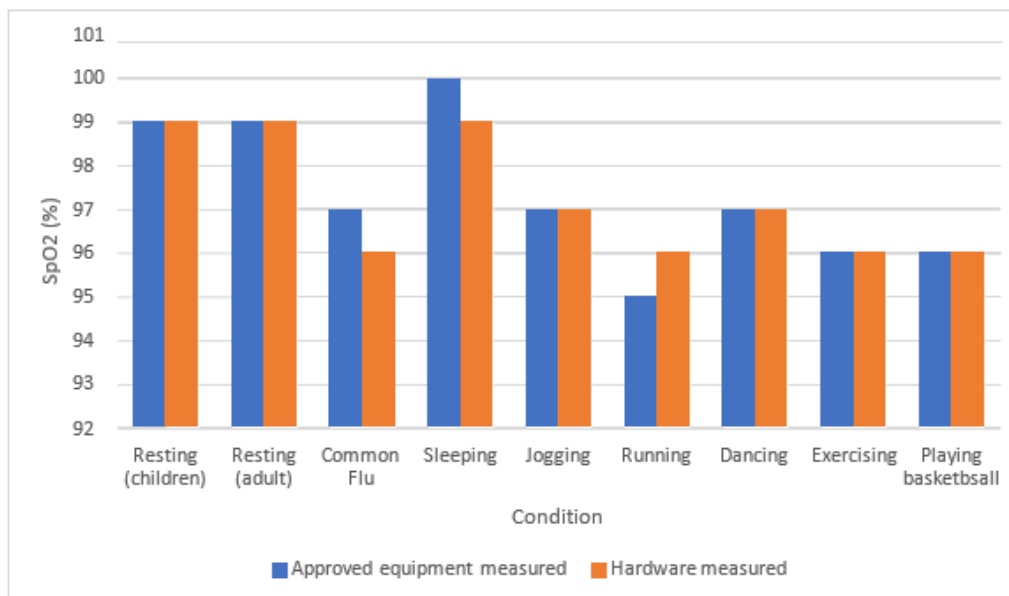


Fig. 13. The Spo2 Taken by Approved Equipment and Hardware.

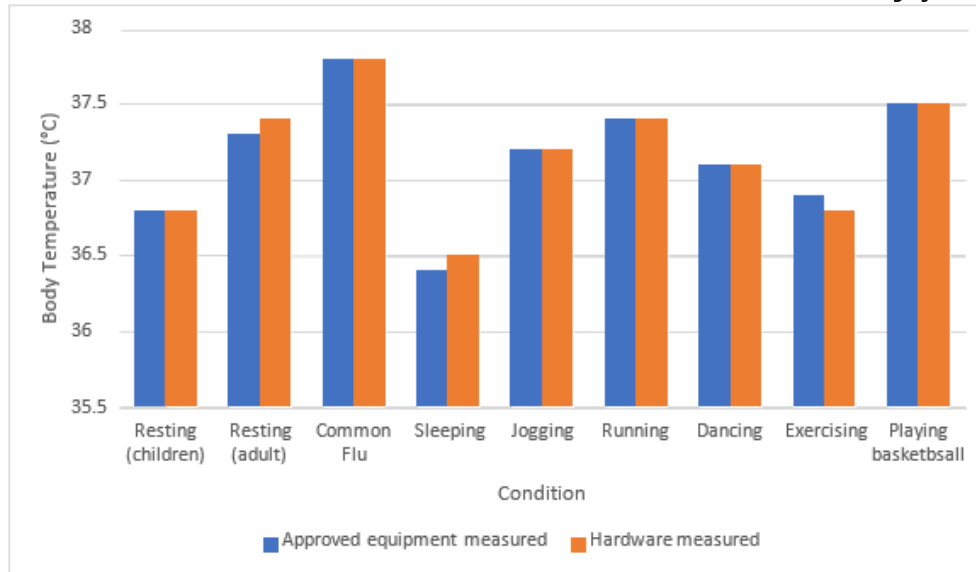


Fig. 14. The Body Temperature Taken by Approved Equipment and Hardware.

Data collected shown in Table I and Fig. 12-14 are the analysis of the measurements taken by the hardware and compare with approved equipment to determine its accuracy. The values were compared with the values measured by approved equipment. In this case, the equipment used to measure the heart rate and SpO2 was a reliable smartwatch meanwhile the body temperature was measured by using the non-contact digital infrared thermometer with stand. The readings measured was varied by the condition of the human as the parameters are closely related to human activities. Heartbeat per minute may increase with vigorous activities as the heart beats faster to provide more blood to the muscles used. The system that can assist in emergency services by enabling the data to be transferred in real time was developed. The system's performance of using NodeMCU as its microprocessor was evaluated. Comparison of the measurements collected from the system of both simulation and hardware with the normal average measurements of the vital signs was made. In the case of hardware, the project was successful in transferring data in real time by using NodeMCU as its microcontroller and Wi-Fi module, bearing the same measurements in both Blynk and OLED display. This system has the ability to improve the existing system of emergency handling situation and emergency handoffs in the emergency department. It benefits each party in the emergency medical services which are the paramedics in the ambulance, doctors, and nurses in the emergency department. The emergency handoffs would become faster and effective. Thus, it improves patient's care in the hospital.

4. Conclusions

In conclusion, overall, the objectives of this project are achieved. The system that can assist in emergency services by enabling the data to be transferred in real time was developed. The system's performance of using NodeMCU as its microprocessor was evaluated. Comparison of the measurements collected from the system of both simulation and hardware with the normal average measurements of the vital signs was made. In the case of hardware, the project was successful in transferring data in real time by using NodeMCU as its microcontroller and Wi-Fi module, bearing the same measurements in both Blynk and OLED display. This system has the ability to improve the existing system of emergency handling situation and emergency handoffs in the emergency department. It benefits each party in the emergency medical services which are the paramedics in the ambulance, doctors, and nurses in the emergency department.

The emergency handoffs would become faster and effective. Thus, it improves patient's care in the hospital.

Acknowledgement

The authors would like to appreciate the involved parties that contributed into successful of this research especially to Universiti Teknologi MARA, Cawangan Pulau Pinang laboratory staffs, the Emergency Department of Hospital Universiti Sains Malaysia (HUSM) and equipment assistant for this project research.

References

- [1] F. Farid, "Conceptual framework of the impact of health technology on healthcare system," *Frontiers in Pharmacology*, vol. 10, no. JULY. 2019. doi: 10.3389/fphar.2019.00933.
- [2] C. Dopfer et al., "COVID-19 related reduction in pediatric emergency healthcare utilization - A concerning trend," *BMC Pediatrics*, vol. 20, no. 1, 2020, doi: 10.1186/s12887-020-02303-6.
- [3] I. B. Lidal, H. H. Holte, and G. E. Vist, "Triage systems for pre-hospital emergency medical services - a systematic review," *Scandinavian Journal of Trauma, Resuscitation and Emergency Medicine*, vol. 21, no. 1. 2013. doi: 10.1186/1757-7241-21-28.
- [4] A. Rahaman, M. M. Islam, M. R. Islam, M. S. Sadi, and S. Nooruddin, "Developing iot based smart health monitoring systems: A review," *Revue d'Intelligence Artificielle*, vol. 33, no. 6. International Information and Engineering Technology Association, pp. 435–440, 2019. doi: 10.18280/ria.330605.
- [5] S. Jayakumar, R. Ranjith Kumar, R. Tejswini, and S. Kavil, "IoT based health monitoring system," *Advances in Parallel Computing*, vol. 39, pp. 193–200, 2021, doi: 10.3233/APC210140.
- [6] M. Taştan, "IoT Based Wearable Smart Health Monitoring System," *Celal Bayar Üniversitesi Fen Bilimleri Dergisi*, pp. 343–350, Sep. 2018, doi: 10.18466/cbayarfb.451076.
- [7] R. Kumar, R. Maheshwari, A. Aggarwal, and M. Shanmugasundaram, "IOT BASED HEALTH MONITORING SYSTEM USING ANDROID APP," vol. 12, no. 19, 2017, [Online]. Available: www.arpnjournals.com
- [8] D. M. P. Sardey, B. Pratik, D. Shalvari, and G. Shubham, "IOT BASED PATIENT HEALTH MONITORING SYSTEM," *International Journal of Advance Engineering and Research Development*, vol. 4, no. 6, 2017.
- [9] R. G. A., S. P., and T. Z. Fadhil, "An efficient IoT based biomedical health monitoring and diagnosing system using myRIO," *Telkomnika (Telecommunication Computing Electronics and Control)*, vol. 18, no. 6, pp. 3050–3057, Dec. 2020, doi: 10.12928/TELKOMNIKA.v18i6.14375.
- [10] Rosylyn Rescue, "Patient Assessment Definitions," Roselyn Rescue Fire Company, 2017.
- [11] S. Bęczkowska, I. Grabarek, S. Pilip, L. Szpakowski, and R. Gała¸zkowski, "Road ambulances: Working conditions of paramedics - Pilot studies," *International Journal of Occupational Medicine and Environmental Health*, vol. 33, no. 1, 2020, doi: 10.13075/ijomeh.1896.01479.
- [12] F. H. Abuzeyad, G. al Qasim, L. Alqasem, and M. I. al Farras, "Evolution of emergency medical services in the Kingdom of Bahrain," *International Journal of Emergency Medicine*, vol. 13, no. 1, 2020, doi: 10.1186/s12245-020-00280-2.

- [13] R. Gerson and V. Feuer, "Innovations in Emergency Assessment and Management of Suicide Risk," *Journal of the American Academy of Child & Adolescent Psychiatry*, vol. 57, no. 10, 2018, doi: 10.1016/j.jaac.2018.07.138.
- [14] S. Lawn, L. Roberts, E. Willis, L. Couzner, L. Mohammadi, and E. Goble, "The effects of emergency medical service work on the psychological, physical, and social well-being of ambulance personnel: A systematic review of qualitative research," *BMC Psychiatry*, vol. 20, no. 1, 2020, doi:10.1186/s12888-020-02752-4.
- [15] D. Davis, "What the ambulance service ought to be," *Journal of Paramedic Practice*, vol. 8, no. 8, 2016, doi: 10.12968/jpar.2016.8.8.394.
- [16] A. Abellsson and L. Lindwall, "What is dignity in prehospital emergency care?," *Nursing Ethics*, vol. 24, no. 3, 2017, doi: 10.1177/0969733015595544.
- [17] N. I. Ismail and N. H. Abdullah, "Malaysia health information exchange: A systematic review," *Business and Economic Horizons*, vol. 13, no. 5, 2017. doi: 10.15208/beh.2017.47.
- [18] M. D. Trust et al., "Prehospital Vital Signs Accurately Predict Initial Emergency Department Vital Signs," *Prehospital and Disaster Medicine*, vol. 35, no. 3, 2020, doi: 10.1017/S1049023X2000028X.
- [19] K. Valentino, G. J. Campos, K. A. Acker, and P. Dolan, "Abnormal Vital Sign Recognition and Provider Notification in the Pediatric Emergency Department," *Journal of Pediatric Health Care*, vol. 34, no. 6, 2020, doi: 10.1016/j.pedhc.2020.05.005.
- [20] A. Kuriyama, S. Urushidani, and T. Nakayama, "Five-level emergency triage systems: Variation in assessment of validity," *Emergency Medicine Journal*, vol. 34, no. 11, 2017. doi: 10.1136/emmermed-2016-206295.
- [21] O. al Shorman, B. al Shorman, M. Al-Khassaweneh, and F. Alkahtani, "A review of internet of medical things (IoMT) - Based remote health monitoring through wearable sensors: A case study for diabetic patients," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 20, no. 1, 2020, doi: 10.11591/ijeecs.v20.i1.pp414-422.

RISK ASSESSMENT TO IMPROVE TOLL ROAD BUSINESS RESILIENCE (CASE STUDY: BALI-MANDARA TOLL ROAD)

Ariyani^{1*},
Andyka Kusuma¹,
Leni Sagita Riantini¹

¹Department of Civil Engineering, Faculty of Engineering, University of Indonesia, Depok, Indonesia

Abstract: Toll roads are one of the transport sectors affected by COVID-19 countermeasure government policies such as large-scale social restrictions and homecoming bans. During the COVID-19 pandemic, the toll road business has its own challenges, which includes a decrease in demand and toll revenues that had been in the operation phase. PT Jasa Marga (Persero) Tbk, as one of the toll road operators in Indonesia, endure the impact of the decline in toll road traffic, which led to a decrease in revenue. According to Jasamarga's annual report, in 2020 when the pandemic began, PT. Jasamarga Bali Tol experienced a 62% decrease in revenue of 88.1 billion rupiah with decrease of 10.4 million vehicles compared to 2019. Revenues associated with toll roads are highly risky due to high initial investment costs, operating and maintenance costs and long concession period. A decline in revenue as a result of the COVID-19 pandemic could disrupt the resilience of the toll road business. This research aims to analyse the risk of declining toll road revenue and its mitigation measures to maintain the resilience of the toll road business itself. This research is carried out by identifying, analysing, evaluating and determining risk treatment options for a decline of toll road revenue. Data collection is performed by literature review and questionnaire survey. Questionnaires were distributed to stakeholders related to the toll road which is the study location. The data collected from the questionnaire results were analysed using descriptive and inferential statistics. This study used 35 risk variables categorised into 8 groups, in which is expected to provide benefits for Concessionaire and related stakeholders in maintaining the resilience of its toll road business.

Keywords: Risk Mitigation, Resilience, Revenue

1. Introduction

1.1 Research Background

At the end of 2019, a new type of virus called SARS-CoV-2 was first detected in the city of Wuhan, Hubei, China. To suppress the spread of the SARS-CoV-2 virus around the world, several countries have implemented lockdown policies. Indonesia was no exception; the government initiated the Large-Scale Social Restrictions policy through the Regulation of the Minister of Health of the Republic of Indonesia Number 9 of 2020. The implementation of PSBB in Indonesia has a significant socio-economic impact. In 2020, economic growth in Indonesia has decreased by 2.07% with the highest decline rate in Central Java at 2.65%. The transport and warehousing sector was the most affected throughout 2020, fluctuating up to 15.04%. Toll Road is one of the transport sectors affected by government policies in tackling COVID-19 such as PSBB and prohibition of homecoming.

During the COVID-19 pandemic, the toll road business were having its own challenges, namely a decline of demand and toll revenues that had been in the operation phase. Toll Road Agency noted that the average daily traffic on toll roads decreased by 80%. The decline of toll road traffic has a negative impact on toll road operators and investors. PT Jasa Marga (Persero) Tbk as one of the toll road operators in Indonesia, endured the impact of the decline in toll road traffic, which was a decline of revenue. Toll road operator PT Jasa Marga (Persero) Tbk estimates that the decline in revenue throughout 2020 due to the COVID-19 pandemic will reach Rp. 2 trillion or approximately 15%. According to Jasamarga's annual report, in 2020 PT. Jasamarga Bali Tol experienced a 62% decrease in revenue of 88.1 billion rupiah with decrease of 10.4 million vehicles compared to 2019. A decrease of revenue due to COVID-19 pandemic can disrupt the resilience of the toll road business. Revenue risk has been shown to be the most important risk associated with toll road provision in the United States (Rouhani, 2015). In India, traffic revenue risk is one of the most critical risks impacting the commercial success of PPP-based transport projects (Singh & Kalidindi, 2006). Research that studies revenue risk on toll roads with BOT project types has been conducted in Indonesia as in the study conducted by (Novianti & Setyawan, 2018). According to (Novianti & Setyawan, 2018), the highest revenue risk is traffic volume risk in excessive fluctuation of traffic projections, either excessively low or excessively high. Mitigation measures in this study are expected to fill the gap in research related to revenue risk in Indonesia. This research aims to analyse the risks that have the potential to reduce toll road revenues and their mitigation measures so that the government and toll road operators can maintain the resilience of their toll road business.

1.2 Literature Review

Risk Management

Uncertainty and risk are two things that are related but not the same. Uncertainty involves complete unknowns, whereas risk can be understood through calculable probabilities. An example often cited to illustrate a risk is a bet on the colour of roulette where the risk of losing is slightly above 50% (Ort & Willumsen, n.d.).

Risk management is a logical and systematic method for identifying, analysing, evaluating, addressing, monitoring, and communicating risks in activities, functions, or processes that enable organisations to minimise losses and maximise opportunities. Risk management is about identifying opportunities as avoiding or reducing losses (Madill, 1933).

Risk management has evolved as an integral part of the PPP procurement process (Le et al., 2020). One important aspect that affects the success of PPP project is risk management (Carbonara et al., 2015). A number of empirical studies on risks in PPP projects have been conducted related to the stages of the risk management process such as risk identification, risk classification, risk analysis, and risk response. The first step of the process is risk identification which helps to recognise potential risk events that affect project implementation (Chan et al., 2011).

According to AS/NZS ISO 31000:2009, there are several stages in risk management illustrated on the scheme below.

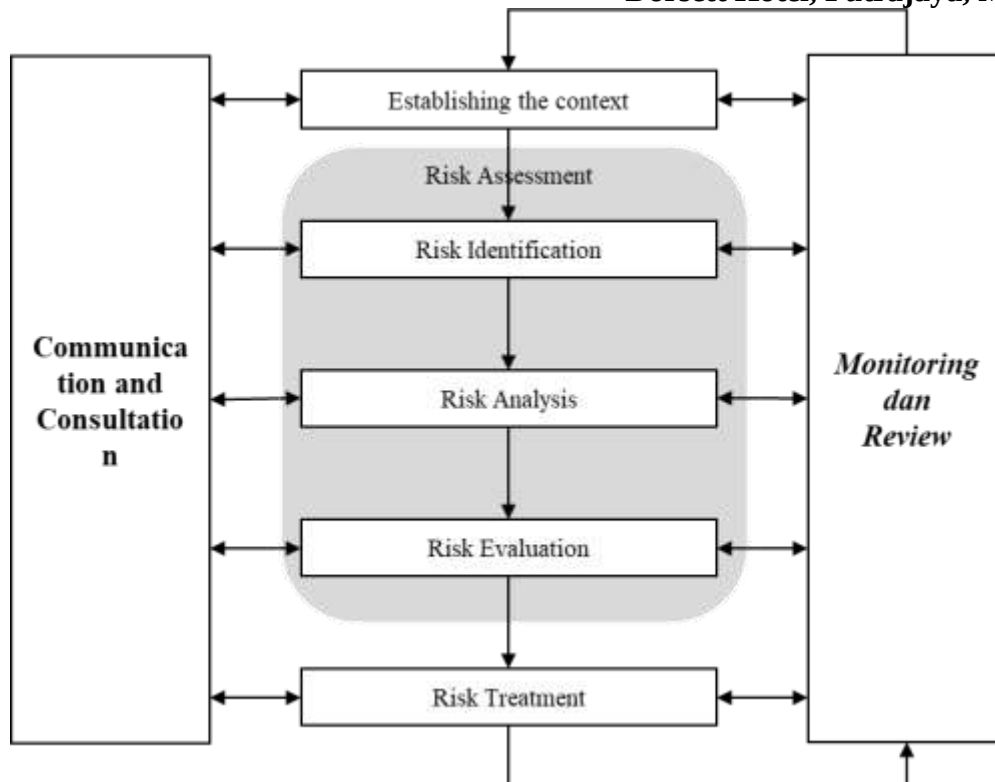


Figure 1: Step of Risk Management (Enterprise Risk Management)

Source: AS/NZS ISO 31000:2009

2. Methodology

This research uses several approaches including literature review, questionnaire survey, quantitative data analysis to determine risk ranking. A comprehensive literature review was used to identify risk variables associated with toll road traffic revenue reduction. The results of the comprehensive literature study resulted in 35 risks of toll road revenue reduction (See table). These risk variables were used to develop a questionnaire as a tool in data collection and the basis for further analysis.

2.1 Samples

The research data was obtained by distributing questionnaires to stakeholders related to the study location toll road. These stakeholders have a role in this study as respondents. According to Baley in Mahmud (2011), the sample size of research using statistical data analysis, has to be at least 30. The research questionnaire was divided into two parts as follows.

- First part was consisted of respondent information including respondent organisation category, academic qualifications, professional experience and other information.
- The second part was structured in relation to the research objectives. The questions asked in the questionnaire had a Likert scale with 5 being set as the highest value.

2.2 Procedures

Several stages carried out in this research are shown in the following diagram.

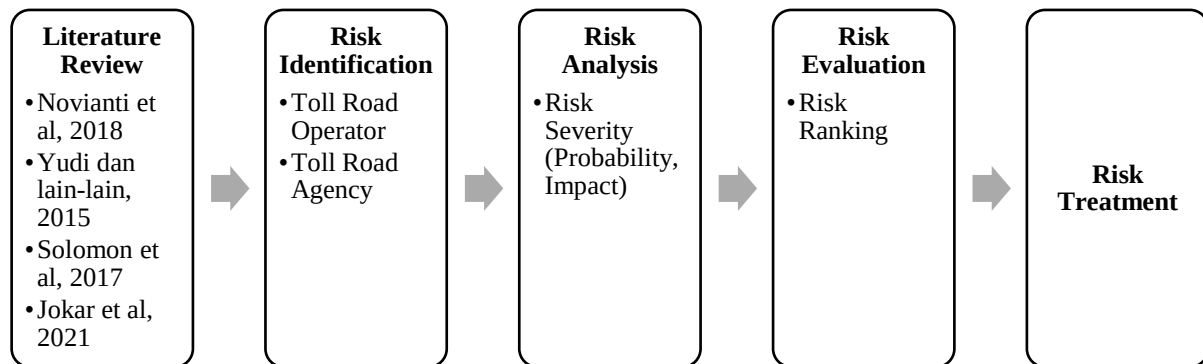


Figure 2: Research Procedure

The first step aims to identify the risks associated with declining toll road traffic revenue. Carbonara et al. (2015) argued that risk analysis and management should be conducted in a specific context. Taking irrelevant risks into the risk register can result in an intricately long questionnaire which potentially impedes the survey rate (Forza, 2002). The interviewees were selected using purposive sampling technique which Saunders et al. explained that it allows the researcher to select cases that will help answer the research questions.

The list of final risk obtained from the experts was then rearranged into a questionnaire to assess risk perception from respondents who have experience working on toll road projects/infrastructure projects. Each respondent provided response regarding the probability and impact on 35 identified risks. The risk level is defined by the multiplication of probability and impact.

Risk evaluation involves comparing the level of risk found during the analysis process with established risk criteria. The output of a risk evaluation is a prioritized list of risks for further action.

Risk treatment involves identifying the range of options for treating risk.

2.3 Data Analysis

2.3.1 Reliability

To determine the suitability of survey-result data collection for analysis, testing, and internal consistency among respondent groups, the Cronbach's alpha coefficient is used. The test results showed that the values for the probability of occurrence and the severity of impact of 35 risk variables were 0.798 and 0.914, respectively. Ideally, Cronbach alpha coefficient of a scale should be above 0.7 (DeVellis, 2012). It should be emphasized, however, that the bigger the number of items in a scale, the higher the Cronbach's (Field, 2018).

2.3.2 Descriptive Analysis

Quantitative data obtained from the questionnaire survey results were processed using descriptive analysis (mean value, standard deviation and frequency analysis). This research uses mean value ranking which is a popular technique in analysing survey data (Chan and Cheung, 2011). The risk score ranking is based on the mean value of the probability of occurrence (P) and severity of impact (I) and the significance of the risk (RS) where higher mean values correspond to higher priorities in the sequence. According to the Project

Management Institute (2017), P and I should be assessed and combined as the core design of the prioritisation process in risk analysis for further investigation and response. According to Chan et al. (2010), Ke et al. (2011), and Shen et al. (2001) risk significance is calculated by multiplying the risk probability and impact as adopted in similar studies such as

$$RS = P \times I$$

(1)

The standard deviation value is used to measure how well the data is presented in average values (Field, 2018). If there is the same mean value on two or more risk variables, the prioritised risk ranking is obtained from the lower standard deviation value (Chileshe and Boadua Yirenkyi-Fianko, 2012). Frequency analysis was used to analyse the demographics of the respondents.

3. Results

3.1 Risk Identification

Based on the results of literature review, previous research, and interviews with toll road operators, the list of risks was identified into 8 categories with 35 sub categories. The risk list is shown in the following table.

Table 1: Risk Identification of Traffic Revenue Decline

Risk Code	Risk Variable	References
A – Toll Rate Risk		
A-1	Delayed government approval of tariff increases during the concession period for legal and political reasons	(Novianti & Setyawan, 2018)
A-2	Low initial tariff proposal due to inaccurate estimation of tariff calculation	(Novianti & Setyawan, 2018), (Harto et al., 2015)
A-3	Enactment of progressive tariffs for overload vehicles that exceed the load limit	(Harto et al., 2015)
A-4	Delay in tariff increase due to non-achievement of toll road minimum service standards	Expert Judgement
B – Traffic Demand Risk		
B-1	Low GDP generation caused by low economic turnover on the projected toll roads	(Novianti & Setyawan, 2018), (Yang et al., 2017)
B-2	Low traffic diversion and traffic flow from existing roads to planned toll roads	(Novianti & Setyawan, 2018)
B-3	Low traffic volume due to low toll road service quality	(Novianti & Setyawan, 2018), (Yang et al., 2017)
B-4	Land use losses caused by failed development of residential or industrial areas around the toll road project	(Novianti & Setyawan, 2018)
B-5	Restriction of traffic volume entering the toll road	(Harto et al., 2015)
B-6	Increased service quality of facilities for other travel mode	(Yang et al., 2017)
B-7	Presence of alternate infrastructure that competes toll roads	(Yang et al., 2017)

Risk Code	Risk Variable	References
B-8	Reduced connectivity to the toll road	(Yang et al., 2017)
B-9	Declining infrastructure growth rate	(Yang et al., 2017)
B-10	Increase of fuel price that affects the cost of vehicle travel	Expert Judgement
B-11	Inaccurate traffic projections (over predict)	Expert Judgement
B-12	Low mobility of toll road users due to the global economic downturn	Expert Judgement
B-13	Low community mobility due to the implementation of the travel warning policy	Expert Judgement
C – Force Majeure Risk		
C-1	Force Majeure caused by socio-political conditions (war, riots, government sanctions)	(Jokar et al., 2021), (Novianti & Setyawan, 2018), (Olusola & Perera, 2017)
C-2	Force majeure caused by natural disasters (flood, landslide, tsunami, earthquake, etc)	(Jokar et al., 2021), (Novianti & Setyawan, 2018), (Olusola & Perera, 2017)
C-3	Force majeure caused by non-natural disaster (Pandemic)	(Jokar et al., 2021), (Novianti & Setyawan, 2018), (Olusola & Perera, 2017)
D – Non Revenue Risk		
D-1	Toll road closure due to construction failure or hidden damage	(Novianti & Setyawan, 2018)
D-2	Free access to toll roads of special considerations	(Novianti & Setyawan, 2018)
D-3	Non-functioning of toll road operator due to decline in project revenue, (corruption)	(Novianti & Setyawan, 2018)
E – Government Risk		
E-1	Government policies restricting infrastructure development, especially toll roads	(Olusola & Perera, 2017), (Yang et al., 2017)
E-2	Faulty funding structure (engaging foreign loans with foreign currency)	(Olusola & Perera, 2017)
E-3	The political resistance to the current government	(Olusola & Perera, 2017)
E-4	Government policies related to changes in transaction systems	Expert Judgement
F – Tolling System Risk		
F-1	Toll road users who pass through the toll road when the system is down	(Olusola & Perera, 2017)
F-2	Damaged toll collection equipment	(Olusola & Perera, 2017)
F-3	Banking system that has a broken transaction system	Expert Judgement
G – Operational and Maintenance Risk		
G-1	Increased operational and maintenance costs	(Jokar et al., 2021)

Risk Code	Risk Variable	References
G-2	Decreased toll road performance and maintenance	(Ng & Loosemore, 2007), (Bing et al., 2005), (Jokar et al., 2021)
G-3	National roads are in better condition than toll roads	Expert Judgement
H – Other Risk		
H-1	Changes in laws and regulations	(Jokar et al., 2021), (Xu et al., 2010), (Kumaraswamy & Zhang, 2001), (Hwang et al., 2013), (Shen et al., 2006), (Bing et al., 2005)
H-2	The inflation rate fluctuations	Expert Judgement

3.1 Risks Analysis

A look-up table or a probability and impact matrix are common methods for determining risk priority (Project Management Institute, 2017). AS/NZS 4360 divided risks into four categories based on P and I values: "Low", "Moderate", "High", and "Extreme". A probability-impact (P-I) matrix was used as an alternate technique to display risk appraisal in this study in order to get a clearer perspective of the priority of risks. Both the X-axis and the Y-axis have values ranging from 1 to 5 ("very low" to "very high"), which are respectively the value of probability and value of impact. Figure below represents the P-I matrix generated based on the survey results.

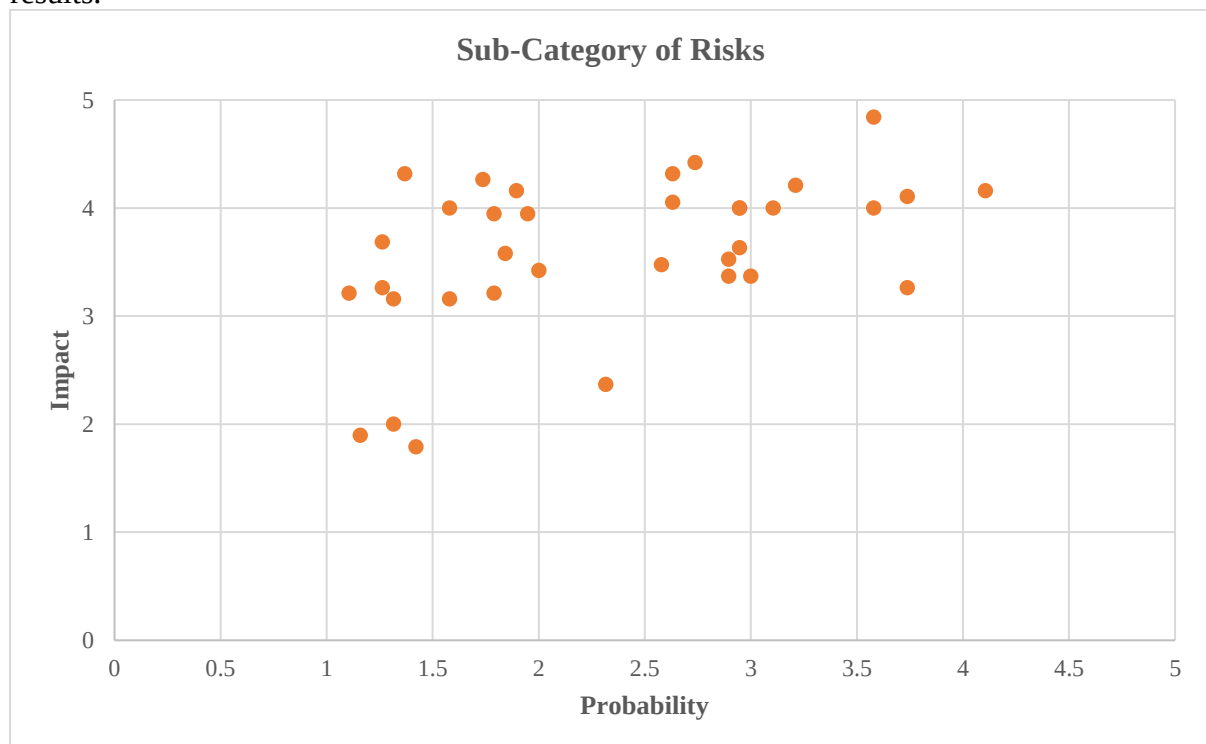


Figure 3: Risk Matrix of Bali – Mandara Toll Road

Table 1: Level of Risk Severity for Bali – Mandara Toll Road

Level of Risk Severity	Risk Category With Different Level of Severity		Remarks
	Amount	Percentage	
Extreme	3	8.57%	Immediate action required
High	18	51.43%	Action required
Moderate	13	37.14%	Mitigation action required
Low	1	2.86%	Accepted

3.3 Risks Evaluation

The outcome of occurrence likelihood analysis and severity of risk impact from the respondents can be seen in table below.

Table 2: Risk Ranking of Bali – Mandara Toll Road

Sub Category of Risk	Mean	Std. Deviation	Level of Risk	Rank
A1	7.0000	1.04013	High	21
A2	7.4737	0.57921	High	19
A3	2.3684	0.49684	Low	35
A4	4.1053	0.45173	Moderate	31
B1	6.3684	0.99475	Moderate	24
B2	15.5263	0.97743	Extreme	3
B3	5.0000	1.19453	Moderate	28
B4	3.3158	0.99723	Moderate	33
B5	9.0526	0.97175	High	17
B6	10.2632	0.80873	High	15
B7	10.7368	0.99072	High	13
B8	11.7368	0.84411	High	11
B9	5.5789	0.57895	Moderate	27
B10	11.9474	0.70700	High	8
B11	14.3684	0.86251	High	4
B12	13.6842	0.96746	High	5
B13	7.8421	0.51449	High	18
C1	6.1579	0.95899	Moderate	25
C2	12.0000	0.44590	High	7
C3	7.3684	0.59339	High	20
D1	10.6842	0.67127	High	14
D2	3.5263	0.35350	Moderate	32
D3	4.3684	0.36842	Moderate	29
E1	11.9474	1.65518	High	9
E2	3.3158	1.08162	Moderate	34
E3	17.1579	0.73454	Extreme	2
E4	4.3158	0.67561	Moderate	30
F1	5.8421	0.63643	Moderate	26

Sub Category of Risk	Mean	Std. Deviation	Level of Risk	Rank
F2	6.5789	0.39968	Moderate	22
F3	10.8421	1.06311	High	12
G1	6.4211	0.6413	Moderate	23
G2	12.3684	1.1678	High	6
G3	9.7895	0.6991	High	16
H1	11.8421	0.7345	High	10
H2	17.2632	0.9454	Extreme	1

Based on table above, 5 most significant risks are the inflation rate fluctuations, the political resistance to the current government, low traffic diversion and traffic flow from existing roads to planned toll roads, inaccurate traffic projections (over predict) and low mobility of toll road users due to the global economic downturn.

3.4 Risk Treatment Option

Toll Road Operators should consider treatment options for dealing with risk of traffic revenue decline. Below are several treatment options can be taken:

- Obtain government guarantees for tariff adjustments or concession extensions
- Maintain road service standards with road maintenance
- Cooperate with the local government in the development of urban areas around toll roads
- Set monetary policy to maintain economic stability
- Obtain traffic/revenue guarantee
- Establish maintenance guarantee
- Conduct market studies and obtain appropriate information on competitive transport mode facilities
- Insure all insurable force majeure risks
- Apply Least Present Value Revenue
- Have a debt reserve account
- Reserve for government revenue shortfall loans
- Using proven technology for toll operations and collection
- Setting aside a budget for unavoidable expenses
- Ensure that inflation has been considered in the tariff calculation in the business plan
- Flexible pricing structure to cover the lack of traffic revenue

4. Conclusions

This paper revealed that the average total score rankings were identified by 35 risk factors for Bali-Mandara toll road revenue decline. Analysis were performed by assessing the probability and impact of each risk subcategory. Risk probability and impact severity indices were determined by combining responses from a questionnaire survey of toll road operator and toll road agency respondents. The study showed the most significant risks are: inflation rate fluctuations, the political resistance to the current government, low traffic diversion and traffic flow from existing roads to planned toll roads, inaccurate traffic projections (over predict) and low mobility of toll road users due to the global economic downturn. Dealing with risk of decline in traffic revenue, toll road operators should consider treatment options: obtain government guarantees for tariff adjustments or concession extensions, maintain road service standards with road maintenance, cooperate with the local government in the development of urban areas around toll roads, set monetary policy to maintain economic stability, obtain traffic/revenue guarantee, etc.

Acknowledgement

This research was not funded by any grant.

References

- [1] Alejandro, Gabriel, and Valdez Cenicerros. n.d. "A Risk Analysis Framework for Toll Road Revenue Forecasts Recommended Citation." https://digitalcommons.utep.edu/open_etd.
- [2] Carbonara, Nunzia, Nicola Costantino, Louis Gunnigan, and Roberta Pellegrino. 2015. "Risk Management in Motorway PPP Projects : Empirical-Based Guidelines Transport Reviews : A Transnational Risk Management in Motorway PPP Projects : Empirical-Based Guidelines," no. October 2017. <https://doi.org/10.1080/01441647.2015.1012696>.
- [3] Chan, Albert P C, John F Y Yeung, Calvin C P Yu, Shou Qing Wang, and Yongjian Ke. 2011. "Empirical Study of Risk Assessment and Allocation of Public-Private Partnership Projects in China," no. July: 136–48. [https://doi.org/10.1061/\(ASCE\)ME.1943-5479.0000049](https://doi.org/10.1061/(ASCE)ME.1943-5479.0000049).
- [4] D'Lima, Minette, and Francesca Medda. 2015. "A New Measure of Resilience: An Application to the London Underground." *Transportation Research Part A: Policy and Practice* 81: 35–46. <https://doi.org/10.1016/j.tra.2015.05.017>.
- [5] Engel, Eduardo, Ronald Fischer, and Alexander Galetovic. 1997. "Infrastructure Franchising and Government Guarantees." *Dealing with Public Risk in Private Infrastructure*, no. March: 89–105. [internal-pdf://174.137.112.147/Dealing with Public Risk in Private Infrastruc.pdf](https://doi.org/10.1016/j.tra.2015.05.017).
- [6] Gauci, Sharon(MITA). 2019. "Enterprise Risk Management Procedure," 15.
- [7] Jin, Jian Gang, Loon Ching Tang, Lijun Sun, and Der Horng Lee. 2014. "Enhancing Metro Network Resilience via Localized Integration with Bus Services." *Transportation Research Part E: Logistics and Transportation Review* 63: 17–30. <https://doi.org/10.1016/j.tre.2014.01.002>.
- [8] Le, Phuong Thi, Nicholas Chileshe, Konstantinos Kirytopoulos, and Raufdeen Rameezdeen. 2020. "Investigating the Significance of Risks in BOT Transportation Projects in Vietnam." *Engineering, Construction and Architectural Management* 27 (6): 1401–25. <https://doi.org/10.1108/ECAM-08-2019-0439>.
- [9] Madill, Ken. 1933. "Standards Association of Australia." *Australian Surveyor* 4 (7): 426. <https://doi.org/10.1080/00050326.1933.10436323>.
- [10] Mattsson, Lars Göran, and Erik Jenelius. 2015. "Vulnerability and Resilience of Transport Systems - A Discussion of Recent Research." *Transportation Research Part A: Policy and Practice* 81: 16–34. <https://doi.org/10.1016/j.tra.2015.06.002>.
- [11] Murray-tuite, Pamela M, and Virginia Tech. 2006. "Proceedings of the 2006 Winter Simulation Conference L. F. Perrone, F. P. Wieland, J. Liu, B. G. Lawson, D. M. Nicol, and R. M. Fujimoto, Eds.," 1398–1405.
- [12] Novianti, T., and H. Y. Setyawan. 2018. "Revenue Risk Modelling and Assessment on BOT Highway Project." *Journal of Physics: Conference Series* 953 (1). <https://doi.org/10.1088/1742-6596/953/1/012145>.
- [13] Olusola, Solomon, and Srinath Perera. 2017. "Analysis of Traffic Revenue Risk Factors in BOT Road Projects in Developing Countries." *Transport Policy* 56 (April 2016): 41–49. <https://doi.org/10.1016/j.tranpol.2017.03.012>.
- [14] Ort, Dios, and Luis G Willumsen. n.d. *MODELLING*.
- [15] Rouhani, Omid M. 2015. "Revenue Risk Mitigation Options for Toll Roads." *IDEAS Working Paper Series from RePEc*, no. 67662.

- [16] Serdar, Mohammad Zaher, Muammer Koç, and Sami G. Al-Ghamdi. 2022. “Urban Transportation Networks Resilience: Indicators, Disturbances, and Assessment Methods.” *Sustainable Cities and Society* 76 (May 2021): 103452. <https://doi.org/10.1016/j.scs.2021.103452>.
- [17] Singh, L Boeing, and Satyanarayana N Kalidindi. 2006. “PROJECT Traffic Revenue Risk Management through Annuity Model of PPP Road Projects in India” 24: 605–13. <https://doi.org/10.1016/j.ijproman.2006.07.008>.
- [18] Truong, Thi My Thanh, and Nga Thi Nguyen. 2022. “Traffic Revenue Risk Identification and Assessment in Build-Operate-Transfer Transport Projects in Vietnam.” *Lecture Notes in Civil Engineering* 203 (January): 1563–72. https://doi.org/10.1007/978-981-16-7160-9_158.

INDUSTRIAL TRIP GENERATION MODEL IN KARAWANG INTERNATIONAL INDUSTRY CITY (KIIC)

Faizal Hendrawan^{1*}
Jachrizal Sumabrata¹
Sutanto Soehodho¹

¹Department of Civil Engineering, Faculty of Engineering, University of Indonesia, Depok, Indonesia

Abstract: The district of Karawang, located in the West Java Province, possesses substantial potential, particularly in the fields of industry, agriculture, and tourism, thereby providing added value to business and investment opportunities. One such example is the Karawang International Industry City (KIIC), an industrial park situated in the district of Karawang. The KIIC spans a total area of approximately 1,347.00 hectares and is home to around 160 companies from around the world, including those from Europe, America, Central Asia, and Southeast Asia. The primary focus of this research is to understand the travel behavior or trip generation and trip distribution of logistics vehicles from various industries within the KIIC. The trip model is analyzed through multiple linear regression using SPSS, which will define the relationship between the number of trips resulting from production or employee commuting. The dominant industries in the KIIC, including automotive, machinery, and electrical industries, are analyzed using variables that influence travel behavior, such as the industrial area's size, travel time, employee income, total production (finished goods), and raw materials. From these variables, it can be determined that employee income and the amount of raw materials significantly influence trip generation in the KIIC.

Keywords: Trip Generation, SPSS, Industry

1. Introduction

Industrial activity is the primary sector that supports the Indonesian economy since the agricultural sector has been replaced as the main economic activity since 1980 (Efendi, T, 2006). Based on the Government Regulation of the Republic of Indonesia Number 14 of 2015 concerning the National Industrial Development Master Plan for 2015 - 2035, it is stated that one of the national industrial development targets is increasing industrial growth which is expected to achieve 2 (two) digit growth in 2035. In fact, as of January 2022 there were 135 industrial estate companies with a total land area of 65,532 Ha spread over the islands of Java, Kalimantan, Maluku, Papua, Nusa Tenggara, Sulawesi and Sumatra with 46% or 30,464 Ha of which were occupied by industrial tenants (Ministry of Industry, 2022).

Java Island is an island with an industrial area that covers almost more than 60% of the area of the national industrial area. Karawang Regency as a Regency in West Java Province with considerable potential, with 12 industrial estates spread across several sub-districts with a total area of around 7,016.98 Ha (Ministry of Industry, 2022). Karawang Regency has an Industrial Area which is located directly adjacent to the center of Karawang City, namely Karawang International Industry City (KIIC) which is 5 km from the city center dominated by the automotive sector industry from Japan.

KIIC which is located in Teluk Jambé District is managed by PT. Maligi Permata Industrial Estate; PT Karawang Tatabina Industrial Estate; PT Harapan Anang Bakri & Sons. The KIIC Industrial Area has a total area of around 1,347.00 Ha which is filled with around 160 companies from around the world such as companies from Europe, America, Central Asia and Southeast Asia (KIIC, 2022). Based on the history of its development, the majority of leasing companies are engaged in the automotive industry only. However, as time goes by, the business fields of tenant companies at KIIC are increasingly diverse, such as the daily product industry and the food industry.

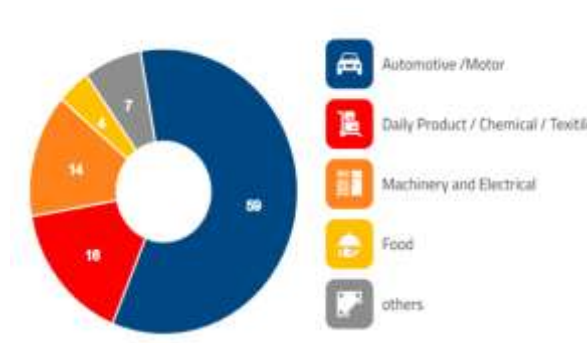


Figure 1. KIIC Industry Ratio

Source: Karawang International Industrial City, 2020

The West Karawang International Road section is the access road to the KIIC industrial area and the West Karawang Toll gate. Logistics vehicles using the toll road from the Jakarta-Cikampek direction can exit directly through the West Karawang 2 toll gate to the KIIC area. Logistics vehicles using the toll road from the Cikampek-Jakarta direction can exit through the West Karawang Toll gate towards the West Karawang International Road section and then turn right towards the KIIC Area.



Figure 2. Accessibility of Karawang International Industrial City (KIIC)

Source: KIIC, 2021

The Karawang Regency Industry and Trade Service reported that industrial company export data reached 7,435 exports in 2019, 4,850 by sea and 350 by air in 2020, and 6,299 exports in 2021. This has caused an increase in logistics vehicle trips and private vehicles in other areas of Karawang Regency, leading to an increase in trip generation and attraction. Additionally, the activity of industrial workers going to KIIC in the morning and returning from KIIC in the afternoon using private vehicles causes an increase in the Level of Service (LOS) on the road (congestion).

Therefore, a more in-depth study of KIIC was carried out to obtain information on the existing location and types of industries in the area, understand the characteristics of trips made from various types of industries in KIIC through logistics vehicle trips, and develop a trip generation model. This research focuses on trip generation and distribution habits of logistics vehicles from various industries in the KIIC area, to facilitate travel trends and estimate the movement of logistics vehicles for new industries.

2. Methodology

2.1 Methodology and Study Region

This research was conducted to analyze the rise and fall of logistics trips in the Karawang KIIC industrial area. The research method starts with the preparatory stage, which identifies problems, targets research objectives, and explores related literature. Next, a logistics vehicle travel survey is conducted to collect travel data from origin to destination. Finally, a trip generation equation is developed and validated to determine the influence of various variables in the KIIC industrial area on trip generation. As in the picture below.

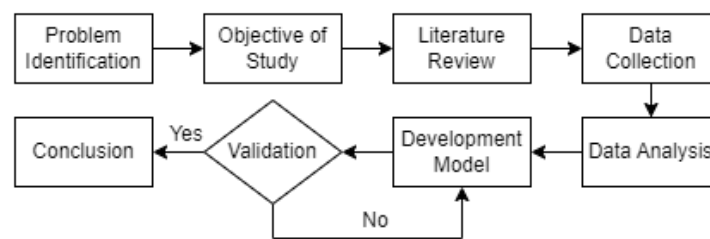


Figure 3. Flow Diagram

This research takes a case study of the Karawang International Industry City (KIIC) industrial area located in Karawang Regency, West Java with an area of 1,500 Ha. It is supported by the environment around Karawang and is close to direct access from the Jakarta-Cikampek toll road. KIIC is occupied by 160 companies from around the world, 75% from Japanese companies. The industrial ratio in KIIC is mostly engaged in the automotive industry, but there are also industries for daily products, machinery and electrical, food, and others. The existence of a large industrial area will have an impact on the industrial area itself even outside the area (between provinces).

Table 1. Classification of Movements Under Study

No	Movement Classification	Description
1	Inter Movement	Inter Movement is Movement from outside to inside KIIC territory or vice versa
2	Inner Movement	Inner Movement is Movement between regions within KIIC



Figure 4. Research Zone

Source: <https://kiic.co.id/id/akses/>, 2021

2.2 Data Collection

The industrial survey was conducted in four industrial categories: automotive/motor, Daily Product/Chemical/Textile, Machinery and Electrical and Food. The results showed that PT Toyota is dominated by 2,702 employee trips in the automotive sector.

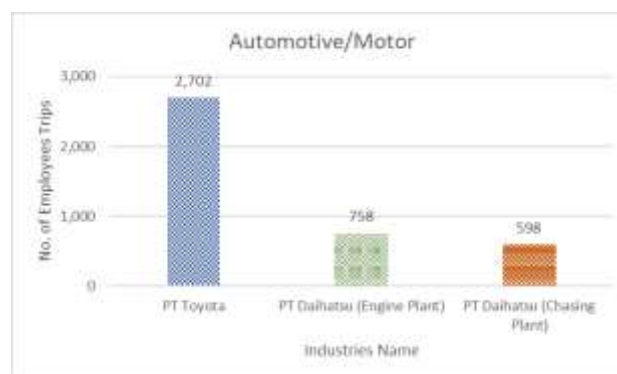


Figure 5. Distribution of Daily Trips to Automotive/Motor Industries

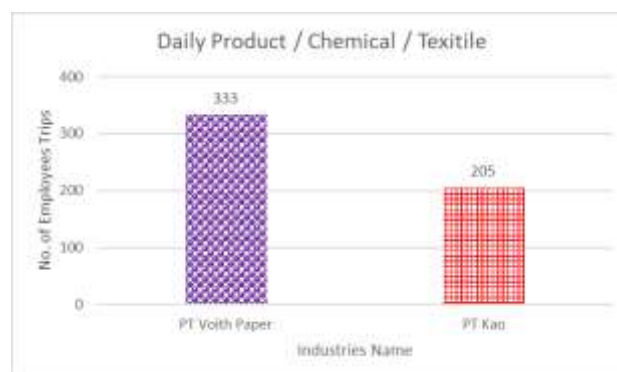


Figure 6. Distribution of daily trips to Daily Product / Chemical / Textile Industries



Figure 7. Distribution of Daily Trips to Machinery and Electrical Industries



Figure 8. Distribution of Daily Trips to Food Industries

The information from industrial goods/freight survey characteristics includes type of total workers, salary, raw materials, finished materials, vehicle type and commodity bulk. Based on this data it will be possible to correlate the number of trips from each industry and the characteristics of the industry and industry area and develop an equation to generate trips.

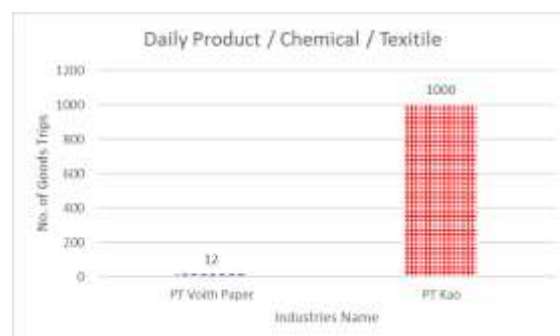


Figure 9. Distribution of Daily Goods Vehicle Trips to Daily Product / Chemical / Textile Industries

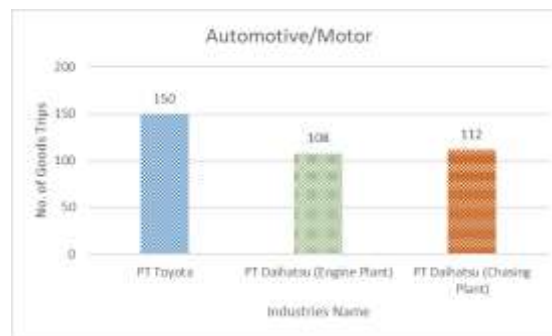


Figure 10. Distribution of Daily Goods Vehicle Trips to Automotive/Motor Industries



Figure 11. Distribution of Daily Goods Vehicle Trips to Machinery and Electrical Industries

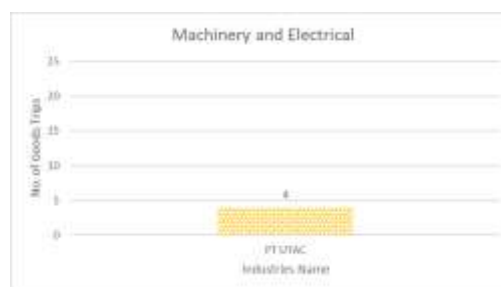


Figure 12. Distribution of Daily Goods Vehicle Trips to Food Industries

The Employees trips survey revealed that motorcycles accounted for 85 percent of total trips per day in the Automotive/motor industries, 51 percent in the Product/Chemical/Textile industries, 34 percent in the Food industries, and 50 percent in the Machinery and Electrical industries. Buses accounted for 70 percent of the total movement in the Food industries, and 50 percent in the Machinery and Electrical industries.



Figure 13. Automotive/Motor and Daily Product/Chemical/Textile Industries Trips/Day (Mode Wise)

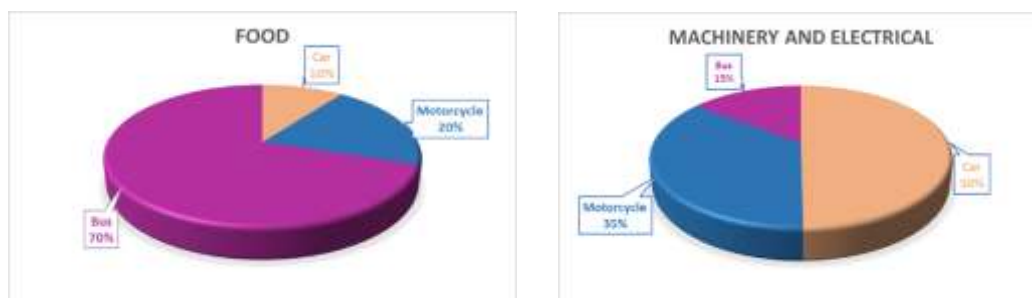


Figure 14. Food and Machinery and Electrical Industries Trips/Day (Mode wise)

3. Results

The multiple linear regression is developed using SPSS, for that number of daily trips is considered as dependent variables and income, sex, employment, trip length, travel time, travel cost, plot area are taken as independent variables. Table 2 summarizes the regression results for the estimated general trip generation model.

Table 2. Industrial Data Base

Company Name	Company Commodity	Num. of Employees (Y)	Area (Ha) (X1)	Ave. Trip time(mi n) (X2)	Ave. Monthly Salary(thousand) (X3)	Ave. Daily Raw material (Ton) (X4)	Ave. Daily Finished material (Unit) (X5)
PT Toyota Motor MFG Indonesia	Automotive	2702	100	27	15000000	150	27000
ADM Engine Plant, PT. Astra Daihatsu Motor -	Automotive	758	16	32	15000000	108	2300

Engine
Plant

Chasing Plant, PT. Astra Daihatsu Motor - Engine Plant	Machiner y	598	16	27	15000000	112	2000
PT Voith Paper Rolls Indonesia	Machiner y	333	5	30	16500000	12	14
PT UTAC Manufactur ing Services Indonesia	Electrical	257	13	22	7250000	4	300000 0
PT. Meiji Food Indonesia	FnB	350	13	28	2000000	75	7
PT Kao Indonesia Chemicals	Chemical	205	12	38	7000000	1000	300

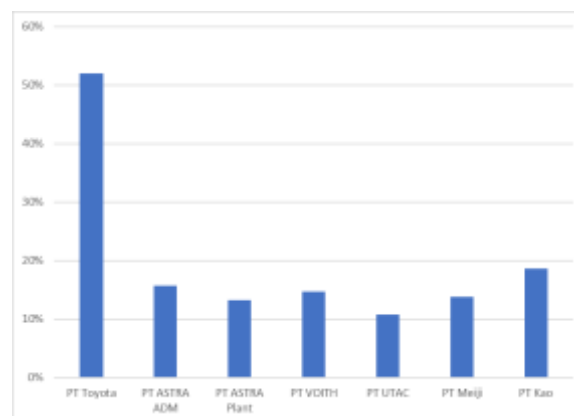


Figure 15. Distribution of Goods Trips to Each Industry

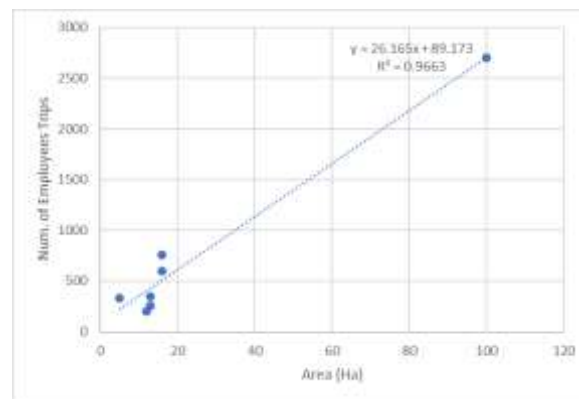


Figure 16. Distribution of Trips as Per Each Industry.

The relationship between the number of daily trips generated per industry and the plot area of the industry is depicted in Figure 16.

Table 3. Results for the general trip generation model

Variables	Coefficient	Standard Error	t-Value	Significance
Constant	-675.167	897.837	-0.752	0.590
Area	25.854	1.960	13.188	0.048
Time	25.074	32.261	0.77	0.579
Salary	1.311E-05	0.000	1.220	0.437
RAW	-0.466	0.358	-1.302	0.417
Finished	-1.626E-05	0.000	-0.204	0.872
R2			0.997	
F-value			68.218	

The final SPSS developed general trip generation model is:

$$Y = -675.167 + 25.854X_1 + 25.074X_2 + 1.311 \times 10^{-5} - 0.466X_4 - 1.626 \times 10^{-5}$$

Where,

Y=Number of daily trips per industry.

X₁= Area (Ha).

X₂=Average Trip Time (min).

X₃=Salary. (Rp)

X₄=Average Daily Raw Material (Ton).

X₅=Average Daily Finished Material.

Validation is the process of comparing simulation results and observational data collected from field studies to determine if the model accurately (Hellings *et al*, 1996) represents the system. It is important to identify the performance measure used in the comparison (Hellings *et al*, 1996). According to (Makridakis, et al., 1982) one measure of forecasting error is the mean

square root or Root mean square error or root mean square deviation. RMSE is a validation method that serves to measure the accuracy of forecast results by finding the average value of the sum of the squared errors. A low value indicates that the results are close to actual conditions.

$$RMSE(\%) = \frac{\sqrt{\frac{1}{n} \sum_{t=1}^n (A_t - F_t)^2}}{\frac{1}{n} \sum_{t=1}^n A_t} \times 100\%$$

Where,

n = lots of data

A_t = actual data

F_t = simulation data

If, RMSE < 5% then the model is acceptable and if RMSE > 10%, it is necessary to review the model, because the model is rejected.

Table 4. General Trip Generation Model Verification

Company Name	Actual Y	Estimated Y	RSME
PT Toyota Motor MFG Indonesia	2702	2713.54	0.00
ADM Engine Plant, PT. Astra Daihatsu Motor - Engine Plant	758	687.15	0.09
Chasing Plant, PT. Astra Daihatsu Motor - Engine Plant	598	559.92	0.06
PT Voith Paper Rolls Indonesia	333	417.05	-0.25
PT UTAC Manufacturing Services Indonesia	257	256.97	0.00
PT. Meiji Food Indonesia	350	354.28	-0.01
PT Kao Indonesia Chemicals	205	213.66	-0.04
Total	5203	6985253.07	1343.54

4. Conclusions

This study developed a trip generation model using linear regression analysis to explain 97% of the variation in daily trips per industry. Seven industrial estates from the KIIC were used, and the model had reasonable explanatory power with an R² value of 0.997. The variables that mostly affect the number of daily trips per industry are area company, consistent with previous studies in developing and developed countries.

Acknowledgement

This research was not funded by any grant.

Copyright © 2023 Global Academic Excellence (M) Sdn Bhd All Rights Reserved

References

- [1] Sahu, P.K., Pani, A. 2020. *A Freight generation and geographical effects: Modelling Freight needs of establishments in developing economies and analyzing their geographical disparities*. Transportation. Vol. 47 Issue 06, pp 2873 – 2902.
- [2] Bhavesh and Chetan. 2021. *Estimating urban freight trips using light commercial vehicles in the Indian textile industry*. Transportation Research Interdisciplinary Perspectives 11 (2021).
- [3] Freight Transport Association. 1996. *Lorries in Urban Areas – Delivering the Goods and Serving the Community*. Freight Matters 5/96, Freight Transport Association.
- [4] Girma dan Emer. 2021. *Modeling Public Transport Users' Trip Production in Hawassa City, Ethiopia*. Journal of Civil Engineering, Science and Technology Volume 12, Issue 2, September 2021.
- [5] Sahu, P.K., Pani, A. 2018. *Modelling Urban Freight generation: A case study of seven cities in Kerala, India*. Transport Policy, 69 (2018) (49 – 64).
- [6] J.Rich. 2015. *Transport Models - From Theory to Practise*. Department of Transport, Technical University of Denmark
- [7] José Holguín-Verasa, Catherine Lawson, Cara Wang, Miguel Jaller, Carlos González-Calderón, Shama Campbella, Lokesh Kalahashtia, Jeffrey Wojtowicz, Diana Ramirez-Ríos. 2017. *Using Commodity Flow Survey Microdata and Other Establishment Data to Estimate the Generation of Freight, Freight Trips, and Service Trips: Guidebook* (2016).
- [8] Alho, A.R., de Abreu e Silva, J. 2016. *Modeling retail Establishments' freight trip generation: A comparison of methodologies to predict total weekly deliveries*. Transportation. Vol. 44 Issue 05, pp 1195 – 1212.
- [9] Meyburg dan Stopher. 1974. *A Framework for the Analysis of Demand for Urban Goods Movement*. Transportation Research Record, (496) (68 – 79).
- [10] N. De Bakshi, G. Tiwari, N.B. Bolia. 2020. *Influence of urban form on urban freight trip generation*. Case Studies on Transport Policy. Vol. 08 Issue 01, pp 229-235.
- [11] Patel, B.V., et al., 2017. *Developing Industrial Trip Generation Model for Himatnagar Industrial Area*. International Journal of Engineering Research & Technology (IJERT). Vol. 6 Issue 04, April-2017.
- [12] Bernardo P. Carlos S. 2020. *Freight trip generation modeling and data collection processes in Latin American cities*. Transportation Research Part A: Policy and Practice. Vol. 132, pp 0965 – 8564
- [13] Tomasz Kulpa. 2014. *Freight Truck Trip Generation Modelling at Regional Level*. Procedia – Social and Behavioral Sciences, (111) (197-202)
- [14] Shahnaz, et al., 2021. *Bangkitan dan Tarikan Pergerakan di Kawasan Pendidikan Kota Bandar Lampung*. Jurnal Transportasi Vol. 21 No. 1 April 2021: 37–44
- [15] V.R. Patel, et al., 2018. *Development of Regional Industrial Trip Generation Model Using SPSS*. Indian Journal of Science and Technology, Vol 11(7), DOI: 10.17485/ijst/2018/v11i7/97742, February 2018.

ANALYSIS OF TRANSPORTATION ACCESSIBILITY IN BOGOR CITY

Agnes Grace Patricia Hutabarat^{1*}
R. Jachrizal Sumabrata¹

¹ Faculty of Engineering, University of Indonesia, Depok, Indonesia

Abstract: Accessibility is the ease of doing mobility from one place to another that belongs to a transportation system. Community mobility occurs every day, this is caused by travel to and from cities or community destinations, such as in the city of Bogor. According to research data from the Global Traffic Scorecard 2021, the level of congestion in Bogor City is ranked 5th in Indonesia, and 821st in the world. The high number of vehicles passing through Bogor City as well as the number of commuters, and the traffic problems experienced in Bogor City, of course disrupt the accessibility of transportation which is an important component of a city and is also very important to ensure that people can access services such as health, education and employment. This cannot be achieved properly if access for community mobility is disrupted. Therefore this study aims to analyze indicators of transportation accessibility using the SEM - PLS method. This study provides an overview of the novelty of accessibility indicators that can be used in analyzing transportation accessibility in Bogor City.

Keywords: Accessibility Indicators, Public Transportation, Structural Equation Model Partial Least Square (SEM – PLS)

1. Introduction

1.1 Research Background

Connectivity according to K. J Kansky's Graphic Theory (1969), is the connectedness of an element with other elements, whether in the form of physical, operational, or other indicators that can be connected. In terms of connectedness, this cannot be separated from the accessibility factor. Accessibility has an important part in the connectivity value of a transportation system. Accessibility is a measure of convenience or convenience regarding the way land use locations interact with each other and the "ease" or "difficulty" of these locations can be reached by the transportation network system (Black, 1981). A movement of people or goods must use a certain transportation system, and one of the factors that people consider in choosing a particular transportation system or mode of transportation is the accessibility factor or the ease factor in reaching or carrying out mobility. The people carry out their daily mobility, going to their destinations such as offices or schools, the same is the case with the people in Bogor City. The city of Bogor, which is one of the cities in West Java Province, is one of the capital buffers with the distance between the City of Bogor and DKI Jakarta about 59 km, while the distance between the City of Bogor and the City of Bandung is around 120 km. The location of Bogor City which has strategic potential for development and growth of the economy and services, the national activity center for industry, trade, transportation, communication and tourism. The center of business activity in Bogor City can be said to be very high and also one of the most densely populated areas in Indonesia. Based on data from the results of the 2021 Global Traffic Scorecard research, Bogor City's congestion level is ranked 5th in Indonesia, and 821st in the world. Then based on Jabodetabek (Jakarta, Bogor, Depok, Tangerang, Bekasi)

Commuter Statistics data for 2019, it was found that the number of commuters from Bogor City whose main activities outside Bogor City were 80,325 people consisting of 76% male and 24% female. While the number of commuters from outside the city of Bogor whose main activities are in the city of Bogor, namely 118,721 people, 59.5% male and 40.5% female, the rest are for school activities, namely 34% and courses 1%. The high number of vehicles passing through Bogor City as well as the number of commuters, and the traffic problems experienced in Bogor City, of course disrupt the accessibility of transportation which is an important component of a city and is also very important to ensure that people can access services such as health, education and employment. This cannot be achieved if transportation accessibility is disrupted, resulting in mobility not being able to run properly. Therefore, this study aims to analyze indicators of transportation accessibility, which later can also be used as a reference for other studies.

1.2 Literature Review

Transportation accessibility in various literatures is said to be the ease for people to reach every place of activity by using a particular transportation system (Tahmasbia B, 2019). Then according to Adharvaryu; B., Chopdeb, A & Dashorac L, 2019) accessibility is a measure of access to a public transportation system that is potentially available in various spatially separated locations. Accessibility is not only physical access to opportunities and services, but also considers affordability, reality and security. Transportation accessibility is important for evaluating the efficiency of public transportation in an area as well as the services of the public transportation. The use of public transportation has a major impact in saving costs and energy, reducing air pollution levels, and solving traffic jams. Improving public transportation accessibility is considered the most effective solution to reduce the level of congestion in the city and can also advance the sustainability of a city. The accessibility of public transportation can be divided into three aspects, namely access to transportation infrastructure, accessibility of the transportation system network, and access to activity points.

2. Methodology

The Structural Equation Model – Partial Least Square (SEM – PLS) method is used to deepen or expand existing accessibility indicators, as well as updating existing accessibility indicators. Prior to the SEM-PLS method stage, however, transport accessibility indicators were found through literature reviews from various journals and publications, which were then listed in the form of a questionnaire.

2.1 Samples

A questionnaire that was given to respondents with educational backgrounds or who were employed in the sectors of urban and regional planning and land transportation was used to collect the data. According to Sugiyono (2014), there should be at least 30 respondents to the questionnaire with validity and reliability testing in order for the calculations to be about normal. Then, according to (Mahmud, 2011), the minimum sample size for research including statistical data analysis is 30, and in accordance with (Effendi, 1995), the evaluation will be more closely aligned with the normal curve if there are at least 30 respondents. Gefen (2000) notes that the minimal sample size for SEM-PLS analysis is at least 10 times the number of indicators in the latent variable with the greatest number of indications. As a result, there is a gap in this study because the sample size is just 40 respondents.

2.2 Procedures

Online surveys with questions about transportation accessibility indicators were then utilized in conjunction with a Likert Scale to gauge people's views, attitudes, and opinions of an event and their objectives in relation to a subject. The existing variables in the Likert Scale are converted to indicator variables based on the Likert Scale. Table 1 lists the factors that were employed in this investigation, which were gleaned through literature reviews.

Table 12: Independent Variable

Number.	Variable	Number.	Code	Indicator	Reference
1.	Distance (X1)	1.	X1.1	Distance of Origin and Destination	(Sonny, 2010) ; (Litman, 2022)
		2.	X1.2	Distance of Origin Location to City Center	(Farida, 2013)
		3.	X1.3	Distance to the Nearest Transportation Infrastructure	(Saghapour et al., 2016) ; (Tonny, 2007)
		4.	X1.4	Walking Distance	(Verseckienė et al., 2016) ; (Fathoni & Erli Handayani, 2022)
2.	Occupation (X2)	5.	X2.1	Total Employment	(Sonny, 2010)
3.	Road (X3)	6.	X3.1	Number of Routes / Tracks / Routes	(Sonny, 2010)
		7.	X3.2	Length of Route / Line / Route	(Sonny, 2010)
		8.	X3.3	Road Construction	(Silva et al., 2019)
		9.	X3.4	Road Surface Evenness	(Silva et al., 2019)
		10.	X3.5	Unobstructed Road Width	(Silva et al., 2019)
		11.	X3.6	Road Crossing Facilities	(Silva et al., 2019)
		12.	X3.7	Network of Connected Roads	(Farida, 2013)
4.	Region (X4)	13.	X4.1	Land Use Intensity / Density	(Priyambodo, 2015)
		14.	X4.2	Regional Topography Conditions	(Farida, 2013)
5.	Time (X5)	15.	X5.1	Travel Time	(Priyambodo, 2015) ; (Fatima et al., 2022) ; (Litman, 2022) ; (Tonny, 2007) ; (Järv et al., 2018)

Number.	Variable	Number.	Code	Indicator	Reference
		16.	X5.2	Average Waiting Time	(Saghapour et al., 2016) ; (Saxena et al., 2021) ; (Fatima et al., 2022) ; (Fathoni & Erli Handayani, 2022) ; (Shah & Adhvaryu, 2016)
		17.	X5.3	Waiting time	(Verseckienė et al., 2016) ; (Jiménez-Espada & González-Escobar, 2021)
		18.	X5.4	Calculation of total access time for each valid route at each Service Access Point	(Fathoni & Erli Handayani, 2022)
6.	Cost (X6)	19.	X6.1	Travel expense	(Priyambodo, 2015) ; (Tonny, 2007)
7.	Transport (X7)	20.	X7.1	Transportation Availability	(Farida, 2013)
		21.	X7.2	Location of Public Transportation Facilities	(Farida, 2013)
		22.	X7.3	Number of Public Transportation	(Saghapour et al., 2016)
		23.	X7.4	Number of Transportation Infrastructure (Stops, Stations, etc.)	(Saghapour et al., 2016) ; (Jiménez-Espada & González-Escobar, 2021)
		24.	X7.5	Number of Bus Routes	(Saghapour et al., 2016) ; (Jiménez-Espada & González-Escobar, 2021)
		25.	X7.6	Number of Bus Routes	(Saghapour et al., 2016)
		26.	X7.7	Ease of accessing public transportation tickets	(Verseckienė et al., 2016)
		27.	X7.8	Ease of access to the public transportation system	(Verseckienė et al., 2016)
		28.	X7.9	The mode used when on the move	(Litman, 2022)

Number.	Variable	Number.	Code	Indicator	Reference
8.	Demography	29.	X8.1	Population density	(Fatima et al., 2022)
		30.	X8.2	Population Age	(Litman, 2022)
		31.	X8.3	Population Income	(Litman, 2022)
		32.	X8.4	Job status	(Litman, 2022)
		33.	X8.5	Population Gender	(Litman, 2022)
9.	Speed	(Shah & Adhvaryu, 2016)			
		34.	X9.1	Walking pace	(Shah & Adhvaryu, 2016)
		35.	X9.2	Average Walking Speed	

The level of accessibility for public transportation is determined by a number of factors, including distance and time. These factors include far-near, long-distance, and fast-distance travel, as well as transportation facilities, infrastructure, infrastructure speed, cost, and other factors. Table 2 contains the following information for the dependent variable.

Table 13: Dependent Variable

Number.	Variable	Number.	Code	Indicator	Reference
1	Public Transportation Accessibility (Y1)	1.	Y1.1	Distance from or to public transportation nodes	(Pramana, 2018) ; (Saxena et al., 2021) ; (Litman, 2022) ; (Shah & Adhvaryu, 2016)
		2.	Y1.2	Walk time	(Saghapour et al., 2016) ; (Saxena et al., 2021) ; (Verseckienė et al., 2016) ; (Fatima et al., 2022) ; (Fathoni & Erli Handayani, 2022) ; (Shah & Adhvaryu, 2016)
		3.	Y1.3	Frequency of each Public Transportation	(Saghapour et al., 2016) ; (Verseckienė et al., 2016) ; (Saxena et al., 2021)

The dependent variable is then helpful for determining and answering the query of how accessible public transit is based on the inquiries on the independent variables above. The results of the questionnaire were used to determine the connection between the independent and dependent variables in this study. These variables were then tested for validity and reliability before being analyzed using the Structural Equation Model Partial Least Square (SEM-PLS).

2.1 Data Analysis

Validity and Reliability

Validity based on (Arifin, 2012), The instrument's validity refers to how accurately and realistically it can measure the target variable. Reliability, according to (Azwar, 2011), is derived from the word reliability, which denotes how much a measurement's results can be relied upon in terms of their consistency, steadiness, and stability. Results of measurements can be believed if they are essentially consistent across multiple measurements of the same group of participants. Validity and reliability are terms used to gauge the caliber of research.

On the outer model, various factors are evaluated, including:

- Outer loading is a measure of how significantly a latent variable and an indicator are correlated. A link between an indicator and its latent variables is closer the higher the outer loading. A value of > 0.7 for the outer loading is acceptable.
- Cross loading is the stage in which the outer loading value of an indicator is assessed in relation to both the latent factors that make up the indicator and to other latent variables.
- A measure of how much a latent or construct variable can explain the variation of its indicators is called average variance extracted (AVE). A value of AVE bigger than 0.5 is considered acceptable because it exceeds the minimum limit of 0.5.
- Cronbach Alpha is the lower limit and composite reliability is the upper limit for internal coherence indicator values. The lowest expected value for exploratory research is 0.70 or 0.60. The upper level to prevent indicator redundancy is 0.95.

To evaluate the model's capacity to explain how the effects of the independent variables collectively (and simultaneously) affect the dependent variable, which can be represented by the value of adjusted R-Squared, the coefficient of determination was tested (Ghozali, 2016). The R-Square value is then classified as strong if it is higher than 0.67, moderate if it is higher than 0.33 but lower than 0.67, and weak if it is higher than 0.19 but lower than 0.33, according to Chin (1998).

3. Results

Following multiple iterations, Table 3 presents the Validity and Reliability utilizing SEM-PLS results.

Table 3: Validity and Reliability

	Cronbach's alpha	Composite reliability (rho_c)	Average variance extracted (AVE)
X3	0,691	0,866	0,764
X5	0,719	0,843	0,642
X7	0,842	0,892	0,675
X8	0,725	0,870	0,771
X9	0,900	0,952	0,908
Y1	0,692	0,829	0,619

The data obtained can be regarded to be valid based on the findings because the Composite Reliability value of Cronbach's Alpha is higher than 0.7 or some are close to 7, it can also be classified as reliable. A latent variable can, on average, explain more than half of the variation of its indicators when the AVE value is more than 0.5, meaning it has good convergent validity. As seen in Figure 1, the structural model.

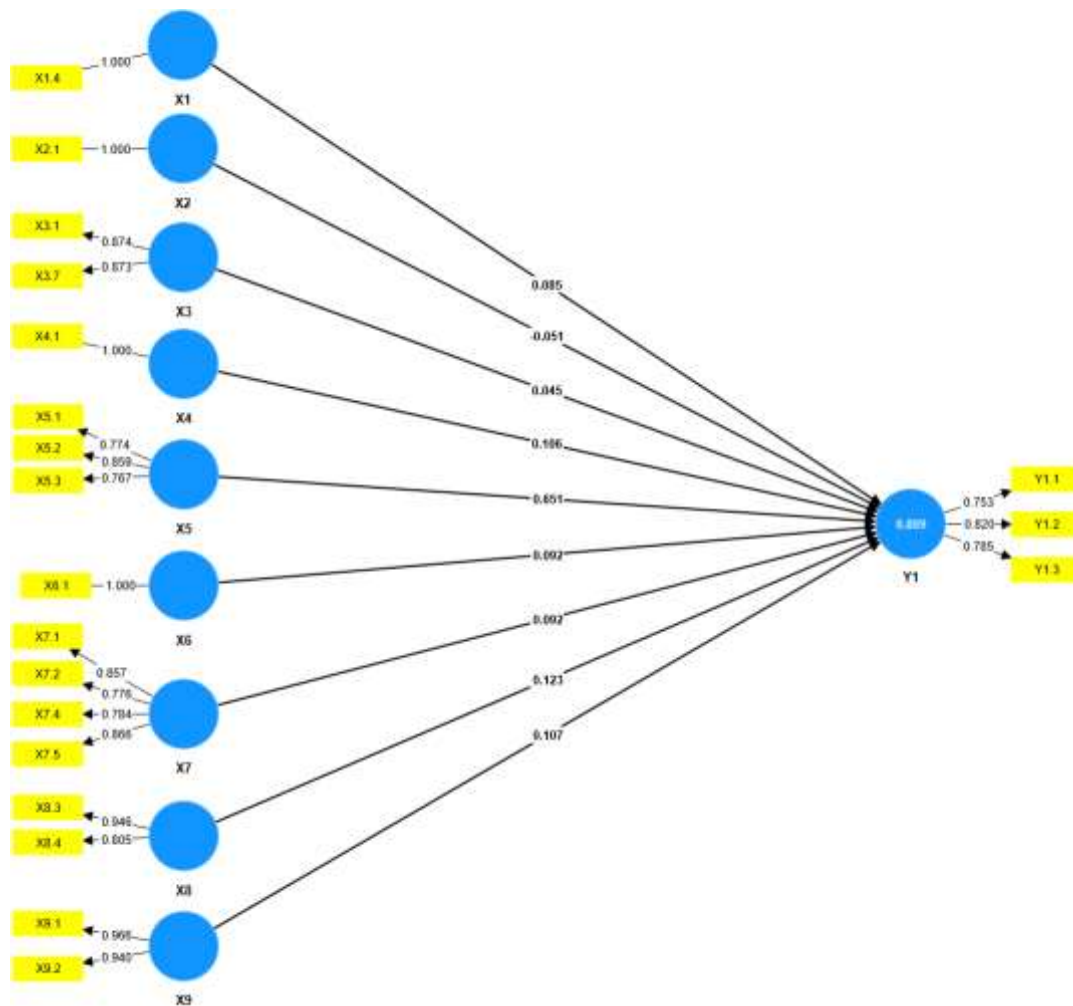


Figure 1: Structural Model

According to Figure 1, the Cross-Loading values are all more than 0.7, indicating a link between the indicator and the construct. After that, for the R Square value of 0.889, which connects Variable X (Table 1) to Variable Y, or the Public Transportation Accessibility Indicator of 88% (Strong) in Table 4. The R-square value is then 85% adjusted for significance. But it may be claimed that Variable X and its value (Table 1) are strongly correlated. Time and Distance are strongly correlated with each other and with the other variables.

Table 4: R square Values

	R-square	R-square adjusted
Y1	0,889	0,856

The result of a T-statistics hypothesis test are shown in Table 5. Variable X is not substantially linked with variable Y if the results are lower than 1.96 (two-tailed), hence the hypothesis is rejected. It is recognized that the hypothesis is true if the P value is less than 0.05. Due to the t-statistics being less than 1.96, all of the hypotheses are disproved as a result.

Table 5: Hypothesis Test

	T Statistics	P Values
X1 -> Y1	0,085	0,085
X2 -> Y1	-0,051	-0,051
X3 -> Y1	0,045	0,045
X4 -> Y1	0,106	0,106
X5 -> Y1	0,651	0,651
X6 -> Y1	0,092	0,092
X7 -> Y1	0,092	0,092
X8 -> Y1	0,123	0,123
X9 -> Y1	0,107	0,107

The research and testing that has been carried out can be ascertained using original and reliable data so that this makes it possible to proceed to the analysis phase using SEM - PLS. Based on the results obtained, there is a correlation between Variable X and Variable Y. However, the T Statistics and P Values are still not correct. It is possible that this happened because it only used 40 samples and is a lack of this study as well.

4. Conclusions

Indicators of accessibility for transportation, including distance, occupation, road, region, time, cost, transportation, demographics, and speed, are analyzed in this study. 40 respondents were surveyed to learn their thoughts and levels of familiarity with the indicators of transportation accessibility. The data in this study are then put to the test using the Validity and Reliability tests. After the SEM-PLS analysis, a significance of 85% between Variable X and Variable Y was discovered. However, because it only has 40 respondents in its sample, it still falls short for P Values and T statistics.

References

- [1] Pramana, A. Y. E. (2018, June 30). Tingkat Aksesibilitas Transportasi Publik di Kota Yogyakarta. REKA RUANG, 1(1), 7–16.
- [2] Saxena, A., Gupta, V., & Shrivastava, B. (2021, June 30). An Assessment of Public Transport Accessibility Levels for Slums in Bhopal. Regular Issue, 10(5), 252–258.
- [3] Litman. (2022, October 19). Evaluating Accessibility for Transport Planning. Victoria Transport Policy Institute.
- [4] Shah, J., & Adhvaryu, B. (2016, September). Public Transport Accessibility Levels for Ahmedabad, India. Journal of Public Transportation, 19(3), 19–35.
- [5] Sonny, I. (2010, October 31). Analisis Indeks Aksesibilitas Transportasi Kabupaten/Kota Di Provinsi Kalimantan Timur. Warta Penelitian Perhubungan, 22(10), 1025–1041
- [6] Farida, U. (2013, April 17). Pengaruh Aksesibilitas Terhadap Karakteristik Sosial Ekonomi Masyarakat Pedesaan Kecamatan Bumijawa Kabupaten Tegal. Jurnal Wilayah Dan Lingkungan, 1(1), 49.
- [7] Saghapour, T., Moridpour, S., & Thompson, R. G. (2016, June). Public transport accessibility in metropolitan areas: A new approach incorporating population density. Journal of Transport Geography, 54, 273–285.
- [8] Tonny. (2007). Measurement of Public Transport Accessibility Condition at Kota Banda Aceh. Jurnal Perencanaan Wilayah Dan Kota, 11(1).

- [9] Verseckienė, A., Meškauskas, V., & Batarlienė, N. (2016). Urban Public Transport Accessibility for People with Movement Disorders: The Case Study of Vilnius. *Procedia Engineering*, 134, 48–56.
- [10] Fathoni, S., & Erli Handayani, K. D. M. (2022, September 12). Pola Spasial Tingkat Aksesibilitas Suroboyo Bus dengan Metode PTAL (Public Transport Accessibility Levels) di Kota Surabaya. *Jurnal Teknik ITS*, 11(2).
- [11] Silva, O. H. D., Pitilin, T. R., Gobbo, C. A. R., Caxambu, M. G., Sanches, S. D. P., & De Angelis Neto, G. (2019, November 29). Accessibility index for urban walkable spaces. *Acta Scientiarum. Technology*, 42, e45181.
- [12] Priyambodo, P. (2015, April 30). Analisis Aksesibilitas dan Level of Service Angkutan Jalan Lintas Surabaya – Kediri. *Warta Penelitian Perhubungan*, 27(2), 129–137.
- [13] Fatima, K., Moridpour, S., & Saghapour, T. (2022, September 12). Development of a Public Transport Accessibility Index for Older Commuters: A Time-Based Approach. *Journal of Advanced Transportation*, 2022, 1–15.
- [14] Järv, O., Tenkanen, H., Salonen, M., Ahas, R., & Toivonen, T. (2018, June). Dynamic cities: Location-based accessibility modelling as a function of time. *Applied Geography*, 95, 101–110.
- [15] Jiménez-Espada, M., & González-Escobar, R. (2021). Research on the problem of universal accessibility in urban public transport. Case study: the city of Cáceres. *Transportation Research Procedia*, 58, 21–28.
- [16] Riawan, A. R., & Ahyudanari, E. (2021, February 19). Analisis Pengaruh Biaya dan Waktu Tempuh Terhadap Aksesibilitas Angkutan Umum Menuju Stasiun KRL, LRT, dan BRT di Kota Bekasi. *Jurnal Penelitian Transportasi Darat*, 22(2), 143–156.

CRYPTOCURRENCY ADOPTION AMONG SAUDI ARABIAN PUBLIC UNIVERSITY STUDENTS: DUAL STRUCTURAL EQUATION MODELLING AND ARTIFICIAL NEURAL NETWORK APPROACH

Ali S. A. Alomari^{1,2*}
Nasuha L. Abdullah¹

¹ School of Computer Sciences, Universiti Sains Malaysia, Pulau Pinang 11800, Malaysia.

² Computer Sciences Department, College of Computing in Al-Qunfudhah, Umm Al-Qura University, Makkah 21955, Saudi Arabia.

Abstract: Cryptocurrency is receiving widespread acceptance in the international market. Unfortunately, little attention is focused on the full identification of the cryptocurrency adoption factors, especially when it comes to emerging nations like Saudi Arabia. The current investigation aims to investigate whether the use of dual Structural Equation Modelling and Artificial Neural Network (SEM-ANN) would permit better comprehension of the determinants of cryptocurrency adoption than the single-step PLS-SEM technique and explore the predictors of cryptocurrency adoption. The Unified Theory of Acceptance and Use of Technology (UTAUT) model was extended with security and awareness. A sample of 344 responses from Saudi Arabian students at public universities was used to verify the model. Unlike the majority of existing studies that were based on a single-step PLS-SEM approach, this investigation employed a superior statistical approach, the dual SEM-ANN approach, considered a unique methodological approach that can recognise variables' connections that are both linear and non-linear and predict relationships with higher accuracy. Moreover, the dual SEM-ANN analysis revealed security as the most important factor influencing whether users would accept cryptocurrency, followed by effort expectancy and awareness. The application of the dual SEM-ANN technique and the extension of the UTAUT model with security and awareness constructs have enhanced the existing technology adoption literature. Additionally, methodological, theoretical, and practical contributions were offered by this study.

Keywords: Cryptocurrency, adoption, determinants, SEM-ANN, Saudi Arabia

1. Introduction

The world has witnessed numerous technological advancements in the last decade, like the internet of things, electronic commerce, and digital payment systems. These changes are reflected in the way individuals interact including how they exchange money [1]. As a result of the continuous change in the means of financial exchange from paper to virtual currency, cryptocurrency has emerged [2]. Cryptocurrency is a digital currency based on a blockchain which is an advanced encryption technique used for financial exchanges. Cryptocurrency transactions are secured, transparent, and anonymous [3]. Moreover, cryptocurrency transactions are not under the control of any financial institutions [4]. Hence, cryptocurrency users enjoy faster transaction time and lower transaction costs.

Furthermore, the recent boom of cryptocurrency has led to a considerable expansion of the cryptocurrency market size [5]. The adoption of cryptocurrency by multinational companies such as Expedia, Wikipedia, Dell and Microsoft, and AliExpress [6] is a sign of a bright future for the technology [7]. The cryptocurrency adoption rate is higher in developed economies than in developing economies like Saudi Arabia, and regulatory policies for cryptocurrency adoption are in the infancy stage [8].

Widespread usage of cryptocurrency was forecasted to help in transforming the national economies, especially in developing nations [6]. Moreover, the potential of cryptocurrency in boosting the e-commerce market has been acknowledged [9]. However, the maximum benefits of cryptocurrency could only be achieved if the technology is widely adopted among individuals [10]. Extensive adoption of cryptocurrency among Saudi Arabians could help the country in achieving its vision 2030 of shifting from an economy based on oil to one based on knowledge [11]. Unfortunately, the adoption of cryptocurrency among individuals in Saudi Arabia is low [12]. Hence, it is crucial to have a thorough understanding of the determinants of cryptocurrency adoption that could stimulate the adoption of the technology among Saudi Arabian users. Previous researchers have called for a deep investigation of the factors of cryptocurrency adoption [13].

The majority of existing studies on cryptocurrency adoption are from the context of developed nations whose cultures and challenges differ from those of Saudi Arabia [13]. The majority of existing studies overlooked Saudi Arabian users' perspectives on cryptocurrency adoption factors [8]. Understanding users' behavioural intention toward cryptocurrency adoption and being able to predict the cryptocurrency adoption factors from the view of Saudi Arabian users may be necessary for the widespread use of cryptocurrency in the country to support the achievement of Saudi Arabian vision 2030 [14].

Additionally, in contrast to other studies utilizing single-stage PLS-SEM [7, 15], that could only identify linear connections among constructs, this investigation utilized dual SEM-ANN that could detect non-linear as well as linear connections between constructs to predict determinants of cryptocurrency adoption more accurately, as recommended by prior studies [16, 17]. This investigation attempts to bridge the literature gaps and achieves the following objectives:

- a) To investigate whether the use of dual SEM-ANN produces more extensive research outcomes in comparison to the single PLS-SEM approach.
- b) To investigate the factors that impact users' decision to utilize cryptocurrency.

Therefore, as recommendations by Wang, Zhang [17] and Lee, Hew [16], this investigation employed a dual SEM-ANN approach to achieve more accurate results. Accordingly, the proposed model of the current investigation can assist cryptocurrency developers with insight for enhancing the features of the technology for its wider adoption among users in emerging economies. This investigation is based on the Unified Theory of Acceptance and Use of Technology model expanded with the insertion of awareness and security.

1.1 Literature Review

Unified Theory of Acceptance and Use of Technology as a suitable model

Some studies examine cryptocurrency adoption through the viewpoint of the Technology Adoption Model [18], however, TAM ignores the influence of social relations and hence is not

deemed suitable for this research [1]. The UTAUT model was acknowledged as a suitable model for the prediction of technology adoption among individuals owing to its high predictive power compared to other theories. UTAUT was developed by [19], through the synthesis of eight different theories, namely: TAM, theory of planned behaviour, theory of reasoned action, model of PC utilization (MPCU), motivational model, combined TAM-theory of planned behaviour, innovation and diffusion theory (IDT), and social cognitive theory. UTAUT model proposes variables as determining factors of behavioural intent to adopt technology, namely: effort expectancy, facilitating conditions, performance expectancy, and social influence. Prior investigations have documented the high reliability of the UTAUT model. However, some scholars have extended the UTAUT model with inclusion of more variables to extend its predictive power. Similarly, we included the constructs of awareness and security to the existing UTAUT constructs for better prediction of cryptocurrency adoption factors from the Saudi Arabian perspective.

1.2 Conceptual Framework of The Study

The constructs employed in the current investigation are based on the unified theory of acceptance and use of technology (UTAUT) and have security and awareness added. The UTAUT constructs of performance expectancy and effort expectancy have both been used extensively in research and are regarded as the most common antecedents of behavioural intent [6]. Additionally, variables such as social influence and facilitating conditions have been rarely studied [6]. Likewise, variables like awareness [20] and security [21] are rarely investigated in the existing literature [6]. Hence, the proposed model of this investigation is on the basis of the UTAUT model extended with awareness and security. These additional constructs are considered crucial in the Saudi Arabian context. Cryptocurrency is an emerging technology. Hence, it is crucial to explore the effect of Saudi Arabians' awareness of cryptocurrency and their perception of security concerning the adoption of the technology as a means of financial exchange. As evident in Figure 1, performance expectancy, effort expectancy, social influence, facilitating conditions, awareness, and security are proposed as determining factor of users' behavioural intent to use cryptocurrency in Saudi Arabia. Detail discussion of hypotheses development is presented in the next section:

1.3 Hypothesis Development and Conceptual Framework

1.3.1 Performance Expectancy (PE)

Performance expectancy may be thought of as the degree to which individuals obtain advantages and values from the adoption of a certain technology like convenience, and ease of use [19]. Performance expectation has been the main indication for influencing users' behavioural intent to use technology in the existing technology adoption literature [22, 23]. It has been demonstrated that before utilizing an innovation, an individual evaluates options using a cost-benefit analysis[24]. Therefore, users' perception of additional benefits of utilizing an innovation would boost their perception of the value of such innovation [19].

Likewise, Rogers, Singhal [25] hints that, an innovation has to provide an additional advantage compared to the existing option for it to be accepted by an individual. The above-mentioned benefits are in line with the performance expectancy. Ho and Ko [26] also reported that users' perception of utility was acknowledged to have a motivating influence on cryptocurrency adoption. As a result, the study anticipates that performance expectancy would enhance the adoption of cryptocurrency among Saudi Arabian users. Consequently, considering the aforementioned arguments, we articulate the hypothesis below:

H1: Performance expectancy has a positive effect on the intention to adopt cryptocurrency.

1.3.2 Effort Expectancy (EE)

Effort expectancy is described as the simplicity with which individuals utilize an innovation [24] that consequently, Impacted the individuals' intention toward the innovation [27]. Studies have revealed the relevance of effort expectancy on intention an individual to adopt a technology [22, 23].

Ter Ji-Xi, Salamzadeh [28] claimed that, apart from the benefits users' get from the adoption of technology, individuals also consider the time and effort required to utilize technology in their decision for adopting the technology. Due to the absence of central control institution, cryptocurrency transactions are fast and easier than other online payment systems such as Paypal. So, it is hoped that effort expectancy would have a positive impact on Saudi Arabian students of public universities to use cryptocurrency. Effort expectancy was acknowledge to be a vital criterion for determining the intention of Saudi Arabians to use cryptocurrency. Many studies have also claimed a critical influence of effort expectancy on users' adoption of cryptocurrency [29]. Therefore, it is anticipated that Saudi Arabian students at public universities would be more inclined to utilize cryptocurrency if it requires less effort to perform financial transactions. Thus, this investigation formulates the hypothesis below:

H2: Effort expectancy has a positive effect on the intention to adopt cryptocurrency.

1.3.3 Facilitating Conditions (FC)

Facilitating condition denotes a person's view that new technology may be used with ease since the necessary technical infrastructure is available [24]. Existing literature revealed empirical evidence implying that an intention of an individual to embrace cryptocurrency is mainly impacted by facilitating conditions [30, 31]. In this research, facilitating conditions reflects the perception of the availability of the necessary resources like internet connectivity, the flexibility of utilizing cryptocurrency on different IT devices, like smartphone, laptop, tablet, and essential knowledge required to conduct a financial transaction using cryptocurrency. Extant literature has revealed the positive influence of facilitating conditions on motivating individuals' to accept cryptocurrency as a means of financial transactions [4, 13, 32]. Therefore, Considering the aforementioned justifications, this investigation proposes that:

H3: Facilitating conditions have a positive effect on the intention to adopt cryptocurrency.

1.3.4 Social Influence (SI)

Social influence refers to individuals' perception of the level to which their friends and family are influencing them to adopt innovation [24]. Social influence has been positively linked to the adoption of several emerging technologies, including mobile banking [27], wearable technologies [33], and mobile wallets [6]. Social influence may be quite important, particularly when there is little knowledge about technology [34]. Cryptocurrency adoption is deemed to be a solitary action. Hence, the positive recommendations on the cryptocurrency benefits from friends and family members could positively influence the students of Saudi Arabian public institutions to use cryptocurrency. The Nseke [35] research, which examined the adoption of cryptocurrency among Africans confirmed this view by offering empirical evidence that showed the stimulating influence of social influence on the cryptocurrency adoption intention.

Many studies reported social influence as a important factor of individuals' intent to adopt cryptocurrency [4, 13, 27, 32]. So, the researchers hypothesized that:

H4: Social influence has a positive effect on the intention to adopt cryptocurrency.

1.3.5 Security

Security refers users' assessment of the level of protection against threats associated with utilizing online technology. People are hesitant to adopt technology when they feel unsecured about using the technology [1]. Cryptocurrency transactions take place virtually. Individuals may be afraid of the possibility of financial loss as a result of cybercrime [36]. The security of cryptocurrency would increase the morale of individuals utilizing the technology, allowing cryptocurrency to gain widespread adoption as a replacement for physical currency [37]. Students at Saudi Arabia's public institutions are more inclined to utilise cryptocurrency if they perceive it to be a secure means of payment [38]. Recent studies have shown that security is a driving force for the adoption of digital currency [29, 39], and [21]. In Malaysia, a lack of security was a significant inhibitor of behavioural intentions to use cryptocurrency [36]. Thus, Saudi Arabian public university students are more probable to utilise cryptocurrencies if they believe it to be a safe technology. The researchers, therefore, postulate that:

H5: Security has a positive effect on behavioural intention to adopt cryptocurrency.

1.3.6 Awareness

Awareness was explained as a person's level of familiarity with innovation and its advantages [40]. We defined awareness as the amount of knowledge Saudi Arabian public university students have concerning cryptocurrency and its advantages. Innovation diffusion theory was the first to suggest the significance of awareness to technological adoption [41]. Cryptocurrency is a new technology. As a result, people lack knowledge of the advantages of cryptocurrency, particularly in emerging nations like the KSA. It has been shown that the level of awareness of an innovation affects the perception of its benefits, and consequently, the behavioural intent to adopt the innovation [41]. There are pieces of evidence indicating that intention of an individual to utilise cryptocurrencies is positively impacted by awareness [40, 42, 43]. However, a lack of user understanding of cryptocurrency might potentially prevent the adoption of cryptocurrency [20]. Thus, further research on the impact of awareness of Saudi Arabian public university students' cryptocurrency adoption behavioural intent. Consequently, we formulated that:

H6: Awareness has a positive effect on behavioural intention to adopt cryptocurrency.

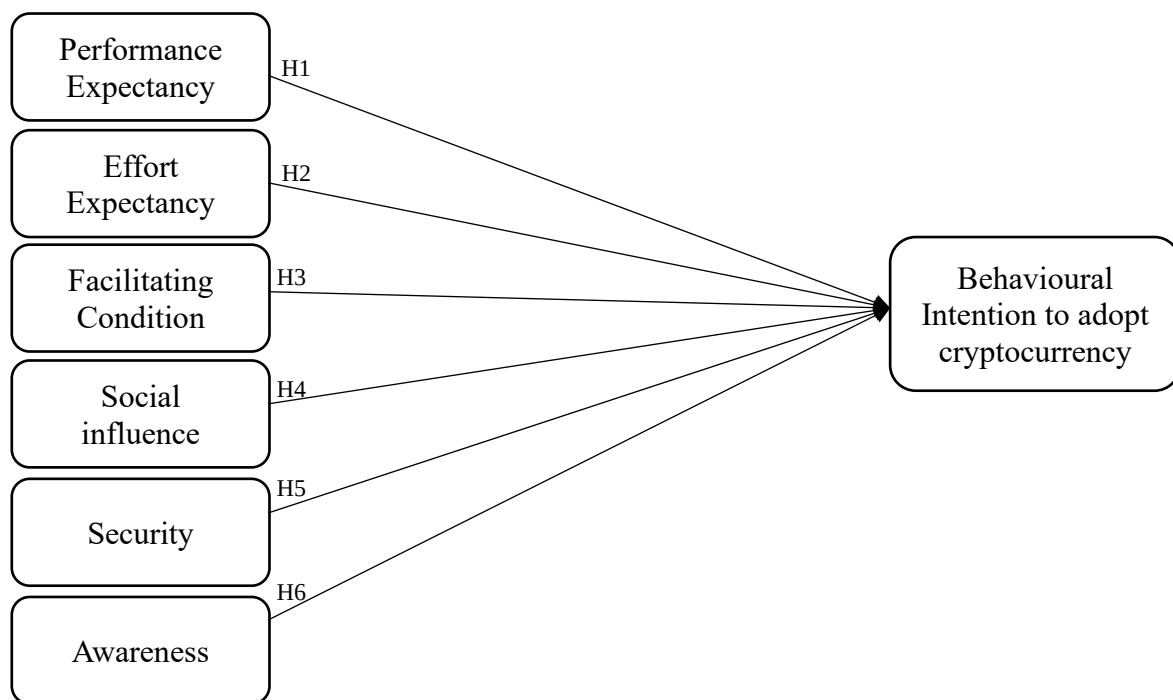


Figure 1: Conceptual Model of The Study

2. Research Methodology

2.1 The Population of The Study

All students at the five largest public universities, specifically: King Fahd University, King Saud University, University of Hail, and Petroleum, Minerals, King Abdulaziz University, and King Khalid University. The five geographical provinces of north, south, east, west, and middle are where the universities are located. The target respondents were selected to be Saudi Arabian public university students because they are thought to be innovators and their use of the technology may encourage other KSA users to adopt it as well [44]. Furthermore, after graduating from college, Saudi Arabian students are projected to play a significant part in the realisation of Saudi Arabia's 2030 goal.

2.2 Data Collection

To fulfil the study's purpose, to gather data, the researchers used online surveys. Bhattacharjee [45] asserts that online surveys have some advantages, i.e. cost-effectiveness, time-saving, and access to a larger audience. We used a non-probability purposive sampling approach to gather data because sampling frame was not accessible. Google survey form was used to create the survey link. The survey link was forwarded to the Deans of student affairs King Fahd University of Petroleum and Minerals, King Abdulaziz University, King Saud University, University of Hail, and King Khalid University. The students were then provided with the survey link by the Deans via their institutions' mailing lists. Data collection took place between August 2022 to October 2022. G*Power was utilized to estimate the minimum sample size, and the results showed that 189 represented the minimum sample size with 0.95 statistical power. 344 valid responses, however, were gathered. All of the variables were measured via a five-point Likert scale. The validity of the survey was pre-tested by academic experts in the field of technology adoption. The survey was pilot tested on a small sample. The reliability of all the constructs was revealed to be more than 0.70, which showed a good pilot test result.

2.3 Data Analysis

In contrast to earlier empirical investigations that used a single-step PLS-SEM analysis technique Sohaib, Hussain [46], the current research has applied a cutting-edge dual SEM–ANN technique to analyse the relationships hypothesised in the conceptual model [16]. There are two phases in this approach. PLS-SEM was first employed to find significant factors of behavioural intent to adopt cryptocurrency. To rank the significant predictors through sensitivity analysis in the next phase, Artificial Neural Network was applied to the significant determinants acquired from the PLS-SEM analysis of the initial stage.

2.4 PLS-SEM

The proposed theoretical model was evaluated via the PLS-SEM analysis. PLS-SEM was preferred over covariance-based SEM (CB-SEM) because it can evaluate complicated models [47], manage small samples, and work with non-normal data [48]. Additionally, it was proposed that PLS-SEM is more effective than CB-SEM [49]. The two components of analysis utilising PLS-SEM are measurement model and structural model assessment [50]. While the outer model analyses validity and reliability, the inner model examines the strength of correlations between variables[48] .

2.5 Artificial Neural Network (ANN).

When PLS-SEM was successfully analysed, an artificial neural network (ANN) is applied to strengthen and verify the PLS-SEM results. Henseler, Ringle [51] assert that the usage of ANN is crucial in predictive studies. ANN is capable of identifying linear and non-linear connections to predict causal relationships with higher accuracy. This study used a multilayer perceptron (MLP) neural network based on feedforward ANN analysis, which has three layers, namely: input, hidden, and output layers. The hidden layer connects the input layer to the output layer. Hidden nodes are represented by weights [52]. The input layer contains the significant factors of the PLS-SEM analysis as its inputs and transmits them in the form of synaptic weights to the hidden layer. The sigmoid function is used for the activation of the hidden layer and output layer [46].

3. Results

3.1 Descriptive Results

Table 1 shows the characteristics of this study's respondents. Generally, 360 responses were gathered, from which 344 were deemed valid for further analysis. Data on demographics were analysed using SPSS 25. Males are the bulk of the respondents (61.3%). By age group, most responders are in the age range of 18 and 24 (54.4%). The majority of responders hold bachelor's degrees (63.1%). Likewise, King Saud University students make up the majority of the respondents (33.7 %).

Table 1: Demographic Profiles of respondents

Demographic Variable	Frequency	Percentage (%)
Gender		
Female	133	38.7
Male	211	61.3
Age Group		
18-24	187	54.4
25-35	104	30.2
36-45	53	15.4
Current Educational Level		
Bachelor's	217	63.1
Master's	92	26.7
PhD	35	10.2
University		
King Fahd University of Petroleum and Minerals	21	6.1
King Saud University	116	33.7
King Abdulaziz University	109	31.7
King Khalid University	68	19.8
University of Hail	30	8.7

Source: Self-created.

3.2 Common Method Bias (CMB)

To account for common method bias, we utilized a combination of statistical and procedural remedies as proposed by Podsakoff, MacKenzie [53]. Harman single factor test was employed. Results show that a single factor's contribution to the overall variance was 34%, which was less than 50% [53]. The researchers also utilized a correlation matrix to further assess common method bias as suggested by [54]. From the correlation matrix result, none of the components had a value of more than 0.90, depicting that CMB was not seen to be a big threat.

3.3 Assessment of the outer model

Measurement model was examined by evaluating the reliability of the constructs through the examination of Cronbach's α , rho_A, and composite reliability (CR). We also examined convergent validity via factor loadings and average variance extracted (AVE), and discriminant validity [49]. Results in Table 2 showed all the factor loadings are more than 0.7 [55], AVE values are higher than 0.5 [56], and Cronbach's α , rho_A, and CR values are all greater than 0.7 [57]. Moreover, discriminant validity was assessed by using the Heterotrait monotrait ratio (HTMT). As disclosed in Table 3, all HTMT ratios are less than 0.85 as proposed by Kline [47]. Hence, No significant issue concerning discriminant validity were discovered.

3.4 Assessment of multicollinearity

This study also examined multicollinearity based on [58]. Multicollinearity could result in an erroneous study. We assess multicollinearity in this study through the evaluation of Variance Inflated Factor (VIF) values. As depicted in Table 4, all VIF values were found to be below 3.3 [29].

Table 2: Measurement Model Evaluation

Construct	Items	Loadings	Cronbach's α	rho_A	CR	AVE
Awareness	AW1	0.824	0.835	0.841	0.902	0.755
	AW2	0.829				
	AW3	0.947				
Behavioural intention	BI1	0.808	0.776	0.777	0.87	0.690
	BI2	0.831				
	BI3	0.853				
Effort Expectancy	EE1	0.790	0.803	0.824	0.871	0.630
	EE2	0.782				
	EE3	0.716				
	EE4	0.878				
Facilitating Condition	FC1	0.777	0.793	0.839	0.876	0.703
	FC2	0.896				
	FC3	0.838				
Performance Expectancy	PE1	0.859	0.852	0.867	0.9	0.693
	PE2	0.760				
	PE3	0.846				
	PE4	0.861				
Security	SEC1	0.894	0.779	0.798	0.872	0.696
	SEC2	0.862				
	SEC3	0.738				
Social Influence	SI1	0.877	0.831	0.835	0.899	0.747
	SI2	0.864				
	SI3	0.852				

as proposed by Kock [59]. demonstrating that the multicollinearity was not a cause for worry in this investigation.

Table 3: HTMT Ratio

Construct	AW	BI	EE	FC	PE	SE	SI
AW							
BI	0.758						
EE	0.766	0.889					
FC	0.127	0.172	0.173				
PE	0.716	0.815	0.823	0.168			
SE	0.688	0.882	0.761	0.157	0.854		
SI	0.679	0.780	0.738	0.090	0.783	0.770	

Note: AW: Awareness; BI: Behavioural intention; EE: Effort expectancy; FC: Facilitating condition; PE: Performance expectancy; SE: Security; SI: Social influence

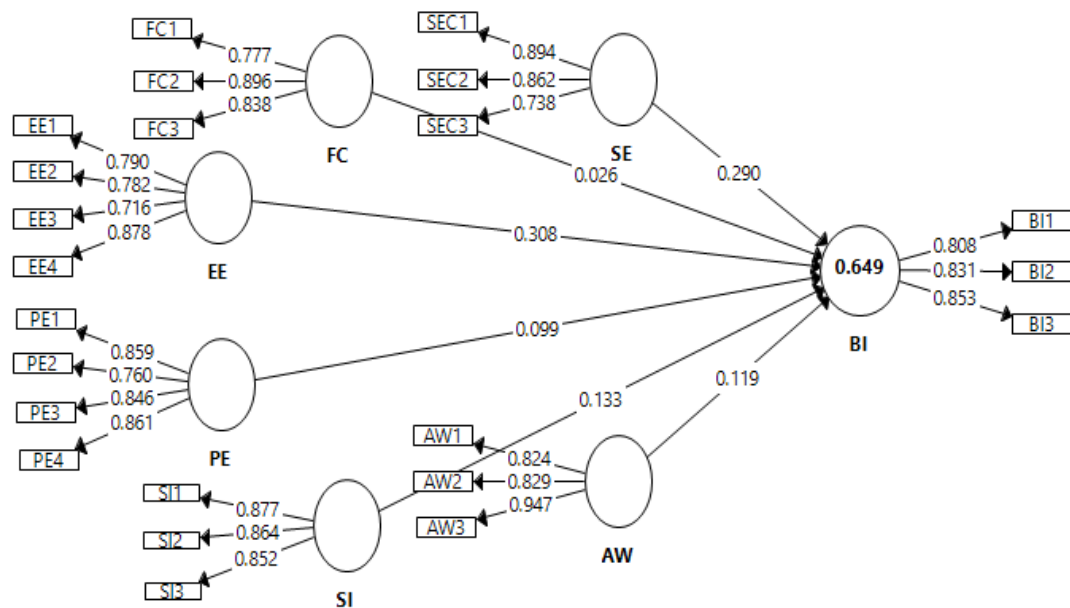


Figure 2: PLS Path Model of The Study

3.5 Assessment of the inner model

When the measurement model was successfully evaluated, the structural model was then assessed. A bootstrapping technique with 5000 iterations was employed to test the proposed hypotheses, as recommended by Hair, Risher [49]. The findings of the hypotheses are shown in Table 4. As shown in figure 2, the combined effect of the proposed determinants accounted for 64.9% of the variance towards behavioural intention to adopt cryptocurrency [60]. The effect size (f^2) of the study constructs and the coefficient of determination (R^2) value are also displayed in Table 4. The blindfolding procedure was also run to determine the predictive relevance (Q^2) of the study model. The Q^2 value was found to be 0.432 as displayed in Table 6. Hence, our path model has a medium predictive accuracy [49].

Table 4: Hypotheses testing

Hypothesis	Relation	β -value	t-value	p-value	Decision	f^2	R^2	Q^2	VIF
H1	AW -> BI	0.119	2.316*	0.021	Supported	0.021	0.649	0.432	1.936
H2	EE -> BI	0.308	5.521*	0.000	Supported	0.114			2.360
H3	FC -> BI	0.290	0.787	0.432	Not Supported	0.006			1.027
H4	PE -> BI	0.099	1.731*	0.084	Supported	0.010			2.800
H5	SE -> BI	0.026	4.640*	0.000	Supported	0.001			2.211
H6	SI -> BI	0.133	2.333*	0.021	Supported	0.021			2.114

Significance level: * $t > 1.645$; $p < 0.05$; ** $t > 2.33$; $p < 0.01$.

Table 4 illustrates that the studied constructs, AW → BI (H1: $\beta = 0.119$, $t = 2.316$, $p = 0.010$), EE → BI (H2: $\beta = 0.308$, $t = 5.521$, $p = 0.000$), PE → BI (H4: $\beta = 0.099$, $t = 1.73$, $p = 0.042$), SE → BI (H5: $\beta = 0.290$, $t = 4.640$, $p = 0.000$), and SI → BI (H6: $\beta = 0.133$, $t = 2.333$, $p = 0.010$), were discovered to have a significant positive influence on users' adoption of cryptocurrency. So, H1, H2, H4, H5, and H6 were supported. However, FC → BI (H3: $\beta = 0.026$, $t = 0.787$, $p = 0.216$) did not have a significant effect. Thus, H3 was not supported.

3.6 Artificial Neural Network Analysis

Artificial neural network analysis was performed through the use of SPSS 25 software, with significant factors of the PLS-SEM analysis (i.e: awareness, security, performance expectancy, effort expectancy, and social influence) as its input (see figure 3).

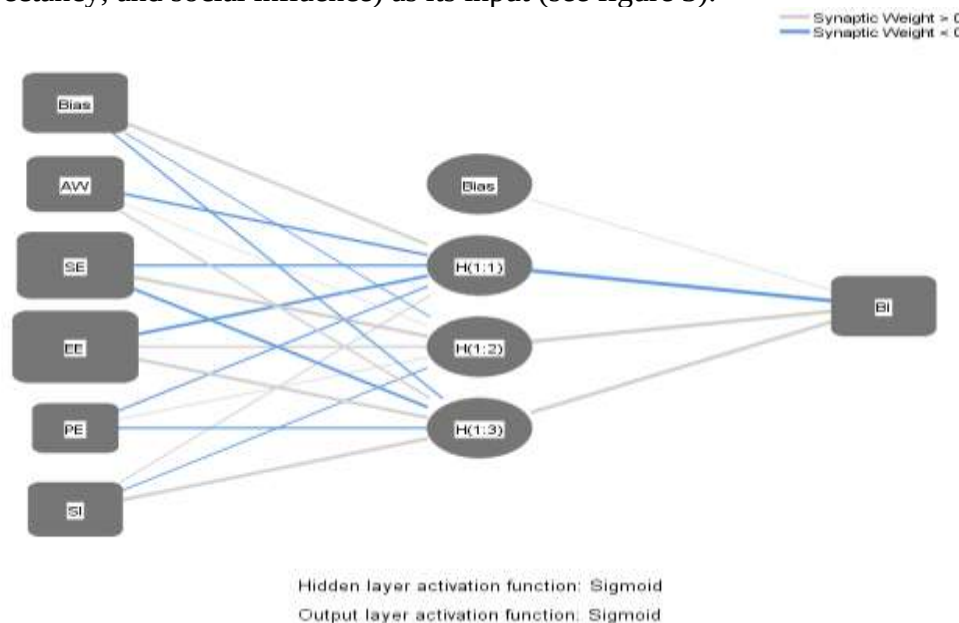


Figure 3: Artificial Neural Network Model of The Study

In this work, both hidden and output neurons are activated using the sigmoid function. To boost the performance of the present model, the range for input and output neurons is also normalised between [0,1] [61]. Additionally, a ten-fold cross-validation approach was used with 10% of the data for testing and 90% for training, to prevent overfitting in ANN models.

To determine the accuracy of our ANN model, the Root Mean Square of Errors (RMSE) was utilized. The RMSE values for training and testing were 0.7883 and 0.821, respectively, as shown in Table 5. The ANN model has great accuracy since the RMSE values for the training and testing data are only slightly different.

Table 5: RMSE Values

Table 3: RMSE Values							
Neural network	Input neurons: SE, AW, EE, PE, SI						Total
	Output neuron: FAD						
	Training			Testing			
	N	SSE	RMSE	N	SSE	RMSE	
1	306	135.135	0.814	38	16.568	0.789	344
2	311	130.207	0.733	33	14.169	0.697	344
3	311	132.283	0.72	33	14.563	0.818	344

4	306	133.354	0.801	38	16.709	0.895	344
5	301	132.636	0.834	43	19.633	0.93	344
6	313	131.18	0.712	31	13.887	0.742	344
7	304	134.666	0.829	38	16.585	0.763	344
8	310	135.521	0.774	34	15.183	0.971	344
9	304	134.129	0.832	40	18.162	0.925	344
10	307	133.992	0.834	37	16.577	0.784	344
Mean		133.31	0.7883		16.2036	0.821	
SD		1.72788	0.04979		1.81073	0.27363	

Notes: SE = Security; AW = Awareness; EE = Effort Expectancy; PE = Performance Expectancy; SI = Social Influence

Furthermore, sensitivity analysis was carried out to determine the normalised importance of the input variables of the ANN model of the current study. To compute the relative normalized importance, each predictor's mean was compared to the highest mean value, represented as a percentage. Details of the normalized importance and the mean importance of each predictor are provided in Table 6. As represented in Table 6, the outcome of the sensitivity analysis disclosed that security is the most important predictor of behavioural intention to adopt cryptocurrency, followed by effort expectancy, and awareness. Moreover, social influence was found to be the least predictor of the behaviour intention of Saudi Arabian public university students to adopt cryptocurrency.

Also, [61] proposed the computation of the goodness of fit of an ANN model as part of the validation of its effectiveness. The goodness of fit is synonymous to the R^2 in the PLS-SEM analysis. To determine the ANN model's predictive power, Equation (1) was used to calculate the goodness-of-fit index.

$$R^2 = 1 - \frac{RMSE}{S_y^2} \quad \text{Equation (1)}$$

Where S_y^2 represents the variance of the preferred output as indicated by the mean (SSE) of the testing process. The outcome suggests that the ANN model predicts 94.93% of cryptocurrency adoption.

Table 6: Sensitivity Analysis with Normalized Importance for Cryptocurrency Adoption

Neural network	Relative importance				
	AW	SE	EE	PE	SI
1	0.167	0.304	0.186	0.158	0.185
2	0.175	0.171	0.258	0.239	0.157
3	0.172	0.189	0.249	0.237	0.153
4	0.163	0.218	0.286	0.171	0.162
5	0.228	0.276	0.184	0.170	0.142
6	0.173	0.198	0.275	0.215	0.139
7	0.210	0.281	0.231	0.119	0.159
8	0.155	0.197	0.275	0.172	0.202

9	0.128	0.240	0.275	0.211	0.146
10	0.169	0.255	0.279	0.112	0.184
Mean relative importance	0.188	0.359	0.264	0.167	0.021
Normalised importance	52.3%	100.0%	73.4%	46.6%	6.0%

Note: AW = Awareness; SE = Security; EE = Effort Expectancy; PE = Performance Expectancy; SI = Social Influence.

The R^2 value of cryptocurrency adoption was 64.9% for the PLS-SEM analysis as opposed to that of the ANN analysis which disclosed an R^2 value of 94.93%. Hence, disclosing a higher predictive power of the ANN model over the PLS-SEM model. Moreover, as implied in Table 7, the results of the PLS-SEM and ANN analysis showed some nominal differences, that are due to the more accurate prediction of the ANN analysis and its ability to detect non-linear relationships among the constructs.

Table 7: Comparison between PLS-SEM and ANN results

Predictors	PLS-SEM		Neural network	
	Path coefficients	Ranking	Relative Importance	Ranking
AW	0.188	3	0.119	4
SE	0.359	1	0.308	1
EE	0.264	2	0.099	5
PE	0.167	4	0.290	2
SI	0.021	5	0.133	3

Notes: AW = Awareness; SE = Security; EE = Effort Expectancy; PE = Performance Expectancy; SI = Social Influence.

4. Discussion

This study aims to explore the determinants of cryptocurrency adoption among students of public universities in Saudi Arabia via the UTAUT model extended with security and awareness. Results of our investigation, all proposed relationships except facilitating condition (H3), awareness, performance expectancy, effort expectancy, security, and social influence (H1, H2, H4, H5, and H6) were found statistically significant predictors of cryptocurrency. The study's findings indicate that users' behavioural intention to use cryptocurrency was positively influenced by awareness (H1: $\beta = 0.119$). This shows that students at Saudi Arabian public universities are expected to have a higher inclination to adopt cryptocurrency when they become familiar with cryptocurrency and knowledgeable about its advantages. Additionally, the outcome of this investigation has firmly established a connection between performance expectancy and cryptocurrency adoption (H4: $\beta = 0.099$). In other words, Saudi Arabian public university students will be more inclined to the adoption of cryptocurrency when they see it as a helpful technology that offers several advantages including time-saving, convenience, and cost-effectiveness. The outcomes of this investigation are in line with prior research [22, 30]. As anticipated, effort expectancy has demonstrated a positive and significant effect on the behavioural intention of users (H2: $\beta = 0.308$), This finding concurs with those of [30]. This hints that making cryptocurrency use simple and convenient is essential to motivating the intention to use cryptocurrency among Saudi Arabian public university students.

The finding from this investigation disclosed that facilitating condition was insignificant to the behavioural intention to adopt cryptocurrency (H3: $\beta = 0.026$). This result is in agreement with those of Shaw and Sergueeva [23] and Merhi, Hone [27]. The possible reason for this outcome could be that, in the context of this study, it is easy to get the prerequisites for using cryptocurrency; all that is needed are a smartphone, tablet, or laptop with internet access to perform cryptocurrency transactions. All these are available to the students of public universities in Saudi Arabia.

A significant impact was found on the association between social influence and behavioural intention to adopt cryptocurrency among students of Saudi Arabian public universities (H6: $\beta = 0.133$). The finding contradicts some existing studies that claim an insignificant impact of social influence on cryptocurrency adoption [21]. Our findings highlight that the opinions of close and loved ones, including family members and peers, affects the intention of Saudi Arabian students of public universities to adopt cryptocurrency.

Security was acknowledged as a crucial predictor of behavioural intention to adopt cryptocurrency among the students of public universities in the KSA (H5: $\beta = 0.290$). The outcome matched that of Almarashdeh, BOUZKRAOUI [62]. This result reflects the impression of worry about the security of financial transactions related to the use of cryptocurrency among students of Saudi Arabian public universities. Security was the most important predictor of behavioural intention to adopt cryptocurrency in this study as revealed by the ANN analysis. Due to the virtual nature of cryptocurrency transactions, security is crucial to the adoption of the technology [38]. This shows that proper safety measures are required to be put in the design of cryptocurrency to prevent any possible risk and hazard associated with cryptocurrency adoption that might deter cryptocurrency usage.

4.1 Contribution to the Theory

The results of the current investigation have advanced the existing information systems literature in general and cryptocurrency in particular by broadening the understanding of the researchers, cryptocurrency developers, Saudi Arabian government and students at public universities in Saudi Arabia regarding the important factors of cryptocurrency adoption. UTAUT was selected as the ideal model since it is recommended to explain the adoption process of emerging innovation [24]. The current study is among the few that have improved the validity to look into and explain the factors that could affect the behavioural intent of students at Saudi Arabian public universities to adopt cryptocurrency as a developing nation. Although the UTAUT model already provides a satisfactory perspective on innovation adoption, it is necessary to incorporate new to propose a model that is capable of interpreting individuals' intentions towards a sensitive technologies such as cryptocurrency [24].

Hence, the current investigation extends the original UTAUT model by incorporating awareness and security constructs. Accordingly, the study findings showed that awareness, performance expectancy, effort expectancy, security, and social influence positively impact the behavioural intention to adopt cryptocurrency, from the view of Saudi Arabian users. The findings of the study enable researchers, cryptocurrency developers, and the Saudi Arabian government to have a thorough grasp of the factors of cryptocurrency in the Saudi Arabian context.

Moreover, unlike prior research that studied the employed a single-stage by utilizing PLS-SEM [23, 27, 46] to predict the complexities associated with the decision to adopt cryptocurrency. This study employs a dual SEM-ANN methodology to determine not only the linear and compensatory links among the research variables, but also the non-linear and non-compensatory correlations among the study variables for more accurate prediction results [63]. The dual SEM-ANN approach is a novel and robust method that is rarely applied in existing cryptocurrency adoption studies [46].

The dual SEM-ANN approach is important since an individual's behaviour intention to adopt cryptocurrency linear relationship of the SEM analysis could not offer a complete explanation of the complex human decision process regarding the adoption of cryptocurrency [64]. Hence this investigation was able to achieve a better prediction outcome than existing studies due to the application of the dual SEM-ANN approach.

4.2 Practical Implications

The empirical results from this investigation also have important practical implications. The findings of this study will help service providers develop strategies that can promote the adoption of cryptocurrency among individuals in developing economies like Saudi Arabia by emphasising ways to make the process of using cryptocurrency more beneficial, simple, and safe for the users. Without widespread user adoption, the implementation of cryptocurrency technology cannot be achieved. To increase the adoption rate in Saudi Arabia, financial institutions and the Saudi Arabian government may raise awareness of the benefits of using cryptocurrencies. Furthermore, the current investigation helps in identifying the important factors affecting cryptocurrency adoption. Thus, providing insight for the Saudi Arabian government and cryptocurrency developers to design effective strategies for enhancing the level of cryptocurrency adoption. The outcomes of the investigation have shown that if users find the technology secure, simple to use and able to help them complete the task, the likelihood of its adoption increases. In addition, Saudi Arabian policymakers must be aware of the significance of the widespread adoption of cryptocurrency to the economic growth of developing nations like Saudi Arabia [5]. Cryptocurrency provides a convenient, low-cost, and fast option for conducting financial transactions [4].

4.3 Limitations and Future Suggestions

Despite its numerous contributions, this research has some limitations. Data collection was performed cross-sectionally but human behaviour changes over time. Therefore, future researchers might consider a longitudinal approach to data collection to better comprehend the influence of progressive changes on the adoption of cryptocurrency among Saudi Arabian users. Secondly, the results of this investigation reflect the perception of Saudi Arabian users and cannot be generalized to other developing countries due to contextual differences. Hence, a similar study is needed in other developing countries to further validate the research model of this study. Thirdly, future investigations might explore the post-adoption behaviour of cryptocurrency users to identify the variables that motivate continuous usage of cryptocurrency among Saudi Arabian users. Lastly, future researchers should evaluate the results of ANNs with more than one layer.

5. Conclusion

The emergence of a new financial payment system by way of cryptocurrency has certainly changed the nature of the global market. This research aimed to explore the factors of

cryptocurrency adoption through an innovative SEM-ANN methodology via the UTAUT model extended with awareness and security constructs. Results showed that awareness, performance expectancy, effort expectancy, security, and social influence are positive factors of cryptocurrency adoption. The outcome enhances the body of knowledge, both theoretically and practically, enabling cryptocurrency developers and policymakers to identify the key factors that could encourage individuals to use cryptocurrency for financial transaction purposes.

Acknowledgement

This research was not funded by any grant

References

- [1] Hutchison, Michael. "Acceptance of electronic monetary exchanges, specifically bitcoin, by information security professionals: a quantitative study using the unified theory of acceptance and use of technology (UTAUT) model." PhD diss., Colorado Technical University, 2017.
- [2] Zhao, Haidong, and Lini Zhang. "Financial literacy or investment experience: which is more influential in cryptocurrency investment?." *International Journal of Bank Marketing* 39, no. 7 (2021): 1208-1226. <https://doi.org/10.1108/ijbm-11-2020-0552>
- [3] Albayati, Hayder, Suk Kyoung Kim, and Jae Jeung Rho. "Accepting financial transactions using blockchain technology and cryptocurrency: A customer perspective approach." *Technology in Society* 62 (2020): 101320. <https://doi.org/10.1016/j.techsoc.2020.101320>
- [4] Omame-Adjepong, Maurice, and Imhotep Paul Alagidede. "High-and low-level chaos in the time and frequency market returns of leading cryptocurrencies and emerging assets." *Chaos, Solitons & Fractals* 132 (2020): 109563. <https://doi.org/10.1016/j.chaos.2019.109563>
- [5] Gil-Alana, Luis Alberiko, Emmanuel Joel Aikins Abakah, and María Fátima Romero Rojo. "Cryptocurrencies and stock market indices. Are they related?." *Research in International Business and Finance* 51 (2020): 101063. <https://doi.org/10.1016/j.ribaf.2019.101063>
- [6] Patil, Pushp, Kuttamani Tamilmani, Nripendra P. Rana, and Vishnupriya Raghavan. "Understanding consumer adoption of mobile payment in India: Extending Meta-UTAUT model with personal innovativeness, anxiety, trust, and grievance redressal." *International Journal of Information Management* 54 (2020): 102144. <https://doi.org/10.1016/j.ijinfomgt.2020.102144>
- [7] Almajali, Dmaithan Abdelkarim, Ra'Ed Masa'Deh, and Zulkhairi Md Dahalin. "Factors influencing the adoption of Cryptocurrency in Jordan: An application of the extended TRA model." *Cogent Social Sciences* 8, no. 1 (2022): 2103901. <https://doi.org/10.1080/23311886.2022.2103901>
- [8] Khwaji, Adel, Farookh Hussain, and Yaser Alsahafi. "Blockchain Technology Adoption in Saudi Hospitals: IT professionals' perspectives." In *Proceedings of the 36th International Business Information Management Association Conference (IBIMA) 4-5 November 2020 Granada, Spain: sustainable economic development and advancing education excellence in the era of global pandemic*. 2021.
- [9] Polasik, Michal, Anna Iwona Piotrowska, Tomasz Piotr Wisniewski, Radoslaw Kotkowski, and Geoffrey Lightfoot. "Price fluctuations and the use of bitcoin: An

- empirical inquiry." *International Journal of Electronic Commerce* 20, no. 1 (2015): 9-49. <https://doi.org/10.1080/10864415.2016.1061413>
- [10] De Filippi, Primavera. "Bitcoin: a regulatory nightmare to a libertarian dream." *Internet Policy Review* 3, no. 2 (2014). <https://doi.org/10.14763/2014.2.286>
- [11] Nurunnabi, Mohammad. "Transformation from an oil-based economy to a knowledge-based economy in Saudi Arabia: the direction of Saudi vision 2030." *Journal of the Knowledge Economy* 8 (2017): 536-564. <https://doi.org/10.1007/s13132-017-0479-8>
- [12] Mukherjee, Subhdeep, Ramji Nagariya, Manish Mohan Baral, Bharat Singh Patel, Venkataiah Chittipaka, K. Srinivasa Rao, and UV Adinarayana Rao. "Blockchain-based circular economy for achieving environmental sustainability in the Indian electronic MSMEs." *Management of Environmental Quality: An International Journal* ahead-of-print (2022). <https://doi.org/10.1108/meq-03-2022-0045>
- [13] Al-Amri, Redhwan, Nur Haryani Zakaria, Adib Habbal, and Suhaidi Hassan. "Cryptocurrency adoption: current stage, opportunities, and open challenges." *International journal of advanced computer research* 9, no. 44 (2019): 293-307. <https://doi.org/10.19101/ijacr.pid43>
- [14] Bhimani, Alnoor, Kjell Hausken, and Sameen Arif. "Do national development factors affect cryptocurrency adoption?." *Technological Forecasting and Social Change* 181 (2022): 121739. <https://doi.org/10.1016/j.techfore.2022.121739>
- [15] Yeong, Yoon-Chow, Khairul Shafee Kalid, K. S. Savita, M. N. Ahmad, and Maryam Zaffar. "Sustainable cryptocurrency adoption assessment among IT enthusiasts and cryptocurrency social communities." *Sustainable energy technologies and assessments* 52 (2022): 102085. <https://doi.org/10.1016/j.seta.2022.102085>
- [16] Lee, Voon-Hsien, Jun-Jie Hew, Lai-Ying Leong, Garry Wei-Han Tan, and Keng-Boon Ooi. "Wearable payment: A deep learning-based dual-stage SEM-ANN analysis." *Expert Systems with Applications* 157 (2020): 113477. <https://doi.org/10.1016/j.eswa.2020.113477>
- [17] Wang, Lizhe, Jiabin Zhang, Peng Liu, Kim-Kwang Raymond Choo, and Fang Huang. "Spectral-spatial multi-feature-based deep learning for hyperspectral remote sensing image classification." *Soft Computing* 21 (2017): 213-221. <https://doi.org/10.1007/s00500-016-2246-3>
- [18] Guych, Nuryyev, Spyridou Anastasia, Yeh Simon, and Achyldurdyeva Jennet. "Factors influencing the intention to use cryptocurrency payments: An examination of blockchain economy." (2018): 303-310.
- [19] Venkatesh, Viswanath, Michael G. Morris, Gordon B. Davis, and Fred D. Davis. "User acceptance of information technology: Toward a unified view." *MIS quarterly* (2003): 425-478. <https://doi.org/10.2307/30036540>
- [20] Mendoza-Tello, Julio C., Higinio Mora, Francisco A. Pujol-López, and Miltiadis D. Lytras. "Social commerce as a driver to enhance trust and intention to use cryptocurrencies for electronic payments." *Ieee Access* 6 (2018): 50737-50751. <https://doi.org/10.1109/access.2018.2869359>
- [21] Hasan, Saher Zeast, Huma Ayub, Abida Ellahi, and Mahnoor Saleem. "A moderated mediation model of factors influencing intention to adopt cryptocurrency among university students." *Human Behavior and Emerging Technologies* 2022 (2022): 1-14. <https://doi.org/10.1155/2022/9718920>
- [22] Alalwan, Ali Abdallah, Yogesh K. Dwivedi, and Nripendra P. Rana. "Factors influencing adoption of mobile banking by Jordanian bank customers: Extending

- UTAUT2 with trust." *International Journal of Information Management* 37, no. 3 (2017): 99-110. <https://doi.org/10.1016/j.ijinfomgt.2017.01.002>
- [23] Shaw, Norman, and Ksenia Sergueeva. "The non-monetary benefits of mobile commerce: Extending UTAUT2 with perceived value." *International journal of information management* 45 (2019): 44-55. <https://doi.org/10.1016/j.ijinfomgt.2018.10.024>
- [24] Venkatesh, Viswanath, James YL Thong, and Xin Xu. "Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology." *MIS quarterly* (2012): 157-178. <https://doi.org/10.2307/41410412>
- [25] Rogers, Everett M., Arvind Singhal, and Margaret M. Quinlan. "Diffusion of innovations 1." *An integrated approach to communication theory and research* (2019): 415-434. <https://doi.org/10.4324/9780203710753-35>
- [26] Ho, Shu-Hsun, and Ying-Yin Ko. "Effects of self-service technology on customer value and customer readiness: The case of Internet banking." *Internet research* 18, no. 4 (2008): 427-446. <https://doi.org/10.1108/10662240810897826>
- [27] Merhi, Mohamed, Kate Hone, and Ali Tarhini. "A cross-cultural study of the intention to use mobile banking between Lebanese and British consumers: Extending UTAUT2 with security, privacy and trust." *Technology in Society* 59 (2019): 101151. <https://doi.org/10.1016/j.techsoc.2019.101151>
- [28] Ter Ji-Xi, Joye, Yashar Salamzadeh, and Ai Ping Teoh. "Behavioral intention to use cryptocurrency in Malaysia: an empirical study." *The Bottom Line* (2021). <https://doi.org/10.1108/bl-08-2020-0053>
- [29] Arias-Oliva, Mario, Jorge Pelegrín-Borondo, and Gustavo Matías-Clavero. "Variables influencing cryptocurrency use: a technology acceptance model in Spain." *Frontiers in psychology* 10 (2019): 475. <https://doi.org/10.3389/fpsyg.2019.00475>
- [30] Beh, Phaik Khee, Yuvaraj Ganesan, Mohammad Iranmanesh, and Behzad Foroughi. "Using smartwatches for fitness and health monitoring: the UTAUT2 combined with threat appraisal as moderators." *Behaviour & Information Technology* 40, no. 3 (2021): 282-299. <https://doi.org/10.1080/0144929x.2019.1685597>
- [31] Gunawan, Fergyanto E., and Rizki Novendra. "An analysis of bitcoin acceptance in Indonesia." *ComTech: Computer, Mathematics and Engineering Applications* 8, no. 4 (2017): 241-247. <https://doi.org/10.21512/comtech.v8i4.3885>
- [32] Mazambani, Last, and Emmanuel Mutambara. "Predicting FinTech innovation adoption in South Africa: the case of cryptocurrency." *African Journal of Economic and Management Studies* 11, no. 1 (2020): 30-50. <https://doi.org/10.1108/ajems-04-2019-0152>
- [33] Lunney, Abbey, Nicole R. Cunningham, and Matthew S. Eastin. "Wearable fitness technology: A structural investigation into acceptance and perceived fitness outcomes." *Computers in Human Behavior* 65 (2016): 114-120. <https://doi.org/10.1016/j.chb.2016.08.007>
- [34] Adapa, Apurva, Fiona Fui-Hoon Nah, Richard H. Hall, Keng Siau, and Samuel N. Smith. "Factors influencing the adoption of smart wearable devices." *International Journal of Human-Computer Interaction* 34, no. 5 (2018): 399-409. <https://doi.org/10.1080/10447318.2017.1357902>
- [35] Nseke, Pissso. "How crypto-currency can decrypt the global digital divide: bitcoins a means for African emergence." *International Journal of Innovation and Economic Development* 3, no. 6 (2018): 61-70. <https://doi.org/10.18775/ijied.1849-7551-7020.2015.36.2005>

- [36] Pandya, Suhag, Murugan Mittapalli, Sri Vallabha Teja Gulla, and Ori Landau. "Cryptocurrency: Adoption efforts and security challenges in different countries." *HOLISTICA–Journal of Business and Public Administration* 10, no. 2 (2019): 167-186. <https://doi.org/10.2478/hjbpa-2019-0024>
- [37] Schaupp, Ludwig Christian, Mackenzie Festa, Kevin G. Knotts, and Elizabeth A. Vitullo. "Regulation as a pathway to individual adoption of cryptocurrency." *Digital Policy, Regulation and Governance* (2022). <https://doi.org/10.1108/dprg-08-2021-0101>
- [38] Al-Mamary, Yaser Hasan, Mohammed A. Alwaheeb, Naif Ghazi M. Alshammari, Mohammed Abdulrab, Hamad Balhareth, and Hela Ben Soltane. "The effect of entrepreneurial orientation on financial and non-financial performance in Saudi SMES: a review." *Journal of Critical Reviews* 7, no. 14 (2020): 270-278. <https://doi.org/10.31838/jcr.07.14.35>
- [39] Mukherjee, Subhdeep, Venkataiah Chittipaka, and Manish Mohan Baral. "Developing a model to highlight the relation of digital trust with privacy and security for the blockchain technology." In *Blockchain Technology and Applications for Digital Marketing*, pp. 110-125. IGI Global, 2021. <https://doi.org/10.4018/978-1-7998-8081-3.ch007>
- [40] Shahzad, Fakhar, GuoYi Xiu, Jian Wang, and Muhammad Shahbaz. "An empirical investigation on the adoption of cryptocurrencies among the people of mainland China." *Technology in Society* 55 (2018): 33-40. <https://doi.org/10.1016/j.techsoc.2018.05.006>
- [41] Ku-Mahamud, Ku Ruhana, Mazni Omar, Nur Azzah Abu Bakar, and Ishola Dada Muraina. "Awareness, trust, and adoption of blockchain technology and cryptocurrency among blockchain communities in Malaysia." *International Journal on Advanced Science, Engineering & Information Technology* 9, no. 4 (2019): 1217-1222. <https://doi.org/10.18517/ijaseit.9.4.6280>
- [42] Bohr, Jeremiah, and Masooda Bashir. "Who uses bitcoin? an exploration of the bitcoin community." In *2014 Twelfth Annual International Conference on Privacy, Security and Trust*, pp. 94-101. IEEE, 2014. <https://doi.org/10.1109/pst.2014.6890928>
- [43] Chang, Younghoon, Siew Fan Wong, Hwansoo Lee, and Seon Phil Jeong. "What motivates chinese consumers to adopt FinTech services: a regulatory focus theory." In *Proceedings of the 18th annual international conference on electronic commerce: e-commerce in smart connected world*, pp. 1-3. 2016. <https://doi.org/10.1145/2971603.2971643>
- [44] Nguyen, Quoc Khanh. "Blockchain-a financial technology for future sustainable development." In *2016 3rd International conference on green technology and sustainable development (GTSD)*, pp. 51-54. IEEE, 2016. <https://doi.org/10.1109/gtsd.2016.22>
- [45] Bhattacharjee, Anol. "Social science research: Principles, methods, and practices." (2012). https://digitalcommons.usf.edu/oa_textbooks/3
- [46] Sohaib, Osama, Walayat Hussain, Muhammad Asif, Muhammad Ahmad, and Manuel Mazzara. "A PLS-SEM neural network approach for understanding cryptocurrency adoption." *Ieee Access* 8 (2019): 13138-13150. <https://doi.org/10.1109/access.2019.2960083>
- [47] Henseler, Jörg, Christian M. Ringle, and Marko Sarstedt. "A new criterion for assessing discriminant validity in variance-based structural equation modeling." *Journal of the*

- academy of marketing science* 43 (2015): 115-135. <https://doi.org/10.1007/s11747-014-0403-8>
- [48] F. Hair Jr, Joe, Marko Sarstedt, Lucas Hopkins, and Volker G. Kuppelwieser. "Partial least squares structural equation modeling (PLS-SEM) An emerging tool in business research." *European business review* 26, no. 2 (2014): 106-121. <https://doi.org/10.1108/eb-10-2013-0128>
- [49] Hair, Joseph F., Jeffrey J. Risher, Marko Sarstedt, and Christian M. Ringle. "When to use and how to report the results of PLS-SEM." *European business review* 31, no. 1 (2019): 2-24. <https://doi.org/10.1108/eb-11-2018-0203>
- [50] Ramayah, T. J. F. H., Jacky Cheah, Francis Chuah, Hiram Ting, and Mumtaz Ali Memon. "Partial least squares structural equation modeling (PLS-SEM) using smartPLS 3.0." *An updated guide and practical guide to statistical analysis* (2018).
- [51] Henseler, Jörg, Christian M. Ringle, and Rudolf R. Sinkovics. "The use of partial least squares path modeling in international marketing." In *New challenges to international marketing*. Emerald Group Publishing Limited, 2009. [https://doi.org/10.1108/s1474-7979\(2009\)0000020014](https://doi.org/10.1108/s1474-7979(2009)0000020014)
- [52] Simpson, Patrick K. *Artificial neural systems: foundations, paradigms, applications, and implementations*. McGraw-Hill, Inc., 1991. <https://doi.org/10.1142/s0129065790000187>
- [53] Podsakoff, Philip M., Scott B. MacKenzie, and Nathan P. Podsakoff. "Sources of method bias in social science research and recommendations on how to control it." *Annual review of psychology* 63 (2012): 539-569. <https://doi.org/10.1146/annurev-psych-120710-100452>
- [54] Lowry, Paul Benjamin, and James Gaskin. "Partial least squares (PLS) structural equation modeling (SEM) for building and testing behavioral causal theory: When to choose it and how to use it." *IEEE transactions on professional communication* 57, no. 2 (2014): 123-146. <https://doi.org/10.1109/tpc.2014.2312452>
- [55] Kavanagh, Donncha, and Paul John Ennis. "Cryptocurrencies and the emergence of blockocracy." *The Information Society* 36, no. 5 (2020): 290-300. <https://doi.org/10.1080/01972243.2020.1795958>
- [56] Loh, Xiu-Ming, Voon-Hsien Lee, Garry Wei-Han Tan, Jun-Jie Hew, and Keng-Boon Ooi. "Towards a cashless society: the imminent role of wearable technology." *Journal of Computer Information Systems* 62, no. 1 (2022): 39-49. <https://doi.org/10.1080/08874417.2019.1688733>
- [57] Glavee-Geo, Richard, Aijaz Ahmed Shaikh, and Heikki Karjaluo. "Mobile banking services adoption in Pakistan: are there gender differences?." *International Journal of bank marketing* (2017). <https://doi.org/10.1108/ijbm-09-2015-0142>
- [58] Kock, Ned, and Gary Lynn. "Lateral collinearity and misleading results in variance-based SEM: An illustration and recommendations." *Journal of the Association for information Systems* 13, no. 7 (2012). <https://doi.org/10.17705/1jais.00302>
- [59] Kock, Ned. "Common method bias in PLS-SEM: A full collinearity assessment approach." *International Journal of e-Collaboration (ijec)* 11, no. 4 (2015): 1-10. <https://doi.org/10.4018/ijec.2015100101>
- [60] Cohen, Jacob, Patricia Cohen, Stephen G. West, and Leona S. Aiken. *Applied multiple regression/correlation analysis for the behavioral sciences*. Routledge, 2013. <https://doi.org/10.4324/9781410606266>
- [61] Liébana-Cabanillas, Francisco, Veljko Marinković, and Zoran Kalinić. "A SEM-neural network approach for predicting antecedents of m-commerce

- acceptance." *International Journal of Information Management* 37, no. 2 (2017): 14-24. <https://doi.org/10.1016/j.ijinfomgt.2016.10.008>
- [62] Almarashdeh, I. B. R. A. H. I. M., H. I. C. H. A. M. Bouzkraoui, A. H. M. E. D. Azouaoui, Hadi Youssef, Lahcen Niharmine, A. A. Rahman, Sharipah Soaad Syed Yahaya, A. M. A. Atta, Dominic Afuro Egbe, And Bethel Mutanga Murimo. "An overview of technology evolution: Investigating the factors influencing non-bitcoins users to adopt bitcoins as online payment transaction method." *Journal of Theoretical and Applied Information Technology* 96, no. 13 (2018): 3984-3993.
- [63] Sternad Zabukovšek, Simona, Zoran Kalinic, Samo Bobek, and Polona Tominc. "SEM-ANN based research of factors' impact on extended use of ERP systems." *Central European Journal of Operations Research* 27 (2019): 703-735. <https://doi.org/10.1007/s10100-018-0592-1>
- [64] Hew, Teck-Soon, Lai-Ying Leong, Keng-Boon Ooi, and Alain Yee-Loong Chong. "Predicting drivers of mobile entertainment adoption: a two-stage SEM-artificial-neural-network analysis." *Journal of Computer Information Systems* 56, no. 4 (2016): 352-370. <https://doi.org/10.1080/08874417.2016.1164497>

CONTRIBUTION OF PLASTIC WASTE IN RECYCLES CONCRETE AGGREGATE PAVING BLOCK

Tommy Iduwin¹
Sigit Pranowo Hadiwardoyo^{1*}
Andri Irfan Rifai²
Riana Herlina Lumingkewas³

¹ Department of Civil Engineering, Faculty of Engineering, Universitas Indonesia, Depok, Indonesia

² Department of Civil Engineering, Faculty of Engineering, Universitas International Batam, Batam, Indonesia

³ Civil Engineering Study Program, Institut Teknologi Indonesia, Tangerang Selatan, Indonesia

Abstract: Road pavements generally use asphalt and cement concrete as flexible and rigid pavements. A flexible pavement covering material with concrete paving blocks has been developed for low-traffic road structures. Paving block compressive strength tests still use beam samples; some have used cube samples. Concrete waste and plastic waste can reduce material requirements and support environmental sustainability. Therefore, it is used as a recycled material. This research utilizes concrete waste and plastic waste in the composition of the concrete paving block mixture. Compressive and flexural strength tests with a cement and aggregate ratio 1:3 was used in this study. The aggregate comprises screening, sand, stone ash, recycled concrete aggregate, and plastic waste additives. The test results showed that the concrete paving block test objects in the form of cubes and beams produced different characteristics. Cube-shaped concrete paving blocks make more realistic compressive strength test values. Judging from the results of the compressive strength and flexural strength tests, the addition of plastic decreased the strength of the concrete paving block. Still, the addition of 4-6% plastic waste positively contributed to the value of flexural strength.

Keywords: Concrete paving block, Compressive strength, Flexural Strength, RCA, Plastic Waste

1. Introduction

Road pavements generally consist of flexible pavements and rigid pavements. Porous and flexible pavements have been developed using paving block surface layers when the vehicle and traffic loads are moderate. Pavement block paving is widely used because of its fast production, easy installation, and maintenance (Yeo et al. 2023). In addition, paving blocks have the ease of working, so they are widely used in road pavement construction. In its application, paving blocks are often used to construct roads, intersections, sidewalks, garages, and parking lots (Djamaluddin et al. 2020). Concrete paving blocks are generally used in areas without vehicle traffic, such as sidewalks, yards, and barriers, or areas with low traffic, such as residential streets, rural roads, and car parks. However, with certain materials, it is also used in areas with medium traffic, such as urban roads, to areas with high traffic, such as industrial areas with a lot of heavy equipment and bus terminals (Mampearachchi 2019).

In the manufacturing process, there are several methods of making paving blocks. The method commonly used is the dry and wet mixed method. Where for the dry method it is made with

very little water with low slump and compacted by a hydraulic machine. In comparison, the wet method is prepared in a mixture with sufficient water and compacted with a vibratory machine (Yeo et al., n.d.). Concrete paving blocks are generally made with cement as a binder, sand as a fine aggregate, and split as a coarse aggregate.

Various mechanical properties tests were carried out to see the performance of the paving block mix. Laboratory tests and field tests on the structural performance of concrete paving blocks show that this pavement system has adequate structural capacity (Jamshidi et al. 2019). The thickness of the concrete paving block affects the structural value of the paving block. Field test results showed that the higher the thickness of the paving block, the lower the deflection value (Arjun Siva Rathan, Sunitha, and Anusudha 2022).

The shape of the paving block also affects the strength value of the paving block. For example, comparing the uni-block and rectangular block shapes of paving blocks shows that the uni-block shape produces a better strength value than the rectangular shape (Bajpai et al. 2023). Several other studies have shown the same tendency that the shape of paving blocks affects the strength value. For example, block I and uni-style shapes produce the same strength values, while clover block shapes produce the lowest strength (Meesaraganda and Kakumani 2021).

Infrastructure development is increasing yearly, indicating the increasing use of materials. This increase has an impact on the limited natural aggregate. Using recycled materials can reduce the use of natural aggregates, which are currently widely programmed in the infrastructure industry, especially in road pavement (Salehi et al. 2021). Sustainability in the construction sector is related to the zero-waste scenario. Its application can be in the form of using recycled materials in construction, one of which is recycled concrete aggregate waste (RCA). Utilization of RCA aggregates can be in the form of mixing in the manufacture of concrete materials, which is beneficial to the environment and financially. RCA has great potential to develop the latest concrete technology, drive economic activity, and optimize natural resources (Makul et al. 2021). The use of RCA in paving block mixes can increase the strength value of concrete. Replacing natural aggregate by 15% will increase strength up to 28% compared to a composition without RCA. This increase is due to adding aggregate and paste bonds (Machado da França and Bianchi Pereira da Costa 2023). In addition, by carrying out proper compaction, paving block composition with RCA content can increase the strength of paving blocks. RCA can increase the strength value of paving blocks by up to 156% against natural aggregate by carrying out proper compaction (Chu et al. 2021). However, when the mixed composition with more dominant RCA will reduce the strength of the mixed composition. Paving blocks with the addition of RCA can reduce strength values by up to 10% after soaking for 24 hours (Luo et al. 2022).

Another waste is plastic waste, a source of problems that are difficult to manage worldwide. In 2050 the accumulation of plastic waste is estimated to reach 12,000 million metric tons (Kankam et al. 2018). The increasing increase in plastic waste makes many people use this plastic waste, including in the construction sector. Therefore, using plastic waste in development is one of the waste management programs to support environmental sustainability (Jawaid et al. 2023).

This research has conducted tests on aggregate and plastic recycled concrete to contribute to environmental sustainability programs and reduce the use of natural materials.

2. Methodology

This study used cement, screening aggregate, rock ash, sand, Recycled Aggregate Concrete, and water. The cement used is a type of cement that is generally used in construction, while aggregate is obtained from PT. Master Block is a paving block manufacturing company. RCA is obtained from the remaining concrete waste from the mixer truck and then processed using a concrete recycling machine. Finally, washing RCA is carried out to remove the mortar content attached to the aggregate.

2.1 Cement

The cement used in the concrete paving block mix is a type of Portland cement used as a binder. Cement material is a type of cement that is commonly used in the manufacture of concrete. The specific gravity value is 3.1 kg/dm^3 in a clean condition without clumping cement particles, so the concrete block mixture has a good bond.

2.2 Natural Aggregate

Natural aggregate consists of screening aggregate material, sand, and rock ash. Screening aggregate has an aggregate size of 12.5 mm to 2.35 mm. While sand with a size of 4.75 mm to 0.075 mm. Stone ash obtained from the results of breaking rock using a stone crusher has an aggregate size of 4.75 mm to 0.075 mm. The results of the screening aggregate properties test, rock ash, and sand can be seen in Table 1.

Table 1: Properties of Agregate

Test type		Standard	SC	AB	Sand	RCA	Unit
Los	Angeles	ASTM C131	36.2	-	-	36.4	%
abrasion							
Specific gravity		ASTM C127	2.43	2.64	2.45	2.3	kg/dm^3
Absorption		ASTM C127	6	2.8	1.5	8.6	%

2.3 Recycled Concrete Aggregate

Recycled Concrete Aggregate (RCA) is obtained from the remaining disposal of the Batching Plant, which accumulates in the stockpile area and contains mortar, so it still needs to be managed. RCA is processed by separating the hardened concrete using a concrete recycling machine. The waste concrete recycling machine produces various sizes of aggregates ranging from coarse, medium, and fine. Furthermore, washing is carried out until the mortar content is lost.



Figure 1: Recycled Concrete Aggregate

2.4 Plastic Waste

Plastic waste of the LDPE type is an added ingredient in a mixture of concrete paving blocks. This waste comes from plastic bags, so processing is required before use. The first process is cleaning the plastic from adhering dirt and then grinding it into a fine and uniform powder. Processing into a fine powder is done by chopping the plastic using a chopping machine. The enumeration results were then filtered using filter No. 16 with a grain size smaller than 1.18 mm.

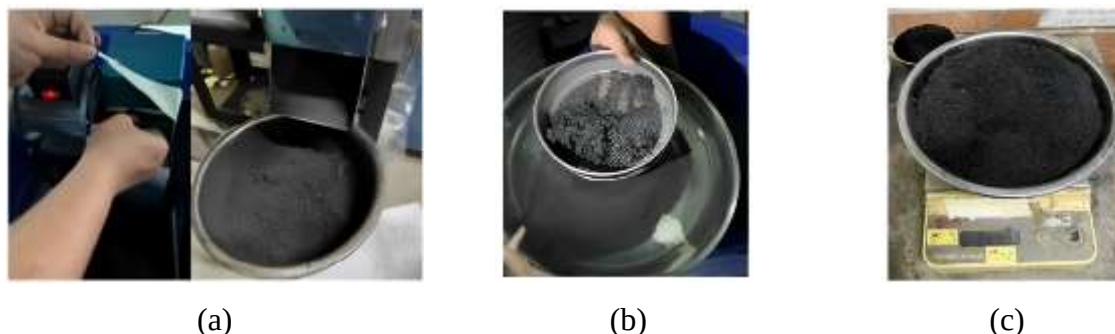


Figure 2: (a) Plastic Chopping Process (B) Plastic Screening Process (C) LDPE Plastic Powder

2.5 Water Cement Ratio

By visual observation, water must be clean and not contain mud, oil, and other floating objects. Water will react directly with the cement so that the cement's hydration takes place, releases heat, and helps the hardening process. The cement water factor is very influential on the compressive strength of concrete. The higher the water content will affect the decrease in compressive strength, and bleeding can occur in the concrete. At this stage, the water requirement will be determined to be used in the paving block mixture, which is 10% of the total mixture.

2.6 Sample Mixtures

The mixture composition used between cement and aggregate is 1:3, with water as much as 10% of the total mixture. There are three types of samples with different heights, namely 10 cm, 8 cm, and 6 cm, with the same width and length of 10 cm and 20 cm. Each weigh 4.13 kg, 3.6 kg, and 2.44 kg. Next, plastic is added to the total weight with a percentage of 4%, 6%, 8%, and 10%. Detailed mixture properties can be seen in Table 2, and a picture of the gradation composition can be seen in Figure 3.

Table 2: Mix Proportion of Concrete Paving Blocks

No	Mix	Water (%)	Cement (%)	Sc (%)	Sand (%)	AB (%)	RCA (%)
1	Control I	10	22	13	27	27	-
2	NA+RCA I	10	22	13	20	20	13
3	Control II	10	22	13	27	27	-
4	NA+RCA II	10	22	13	20	20	13
5	Control III	10	22	13	27	27	-
6	NA+RCA III	10	22	13	20	20	13

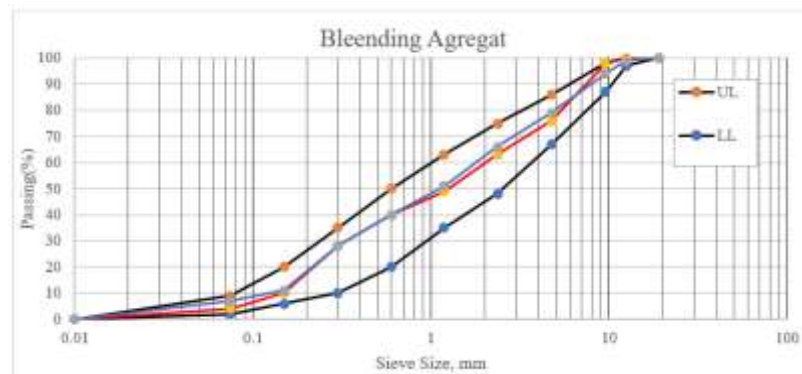


Figure 2: Blending Agregat (a) without RCA (b) Modified RCA

2.7 Sample Preparation

The concrete paving block sample process is done manually by preparing tools and materials. First, the mold is greased with oil so that it does not stick to the mold when removing the test object. The compacting concrete paving blocks using a manual compactor. The compaction process is carried out by dividing the three layers, each being punched 45 times until the mold is filled. Then the mold and compactor were removed and left to stand for up to 3 hours before the test object could be moved. After the specimens have hardened, the specimens are divided into beams and cubes. The cube test object obtained from the sample beam is cut into cubes of 90 mm, 70 mm, and 50 mm.



(a)



(b)



(c)

Figure 3: (a) Beam Sample (B) Cube Sample (C) Process Of Cutting Cube Sample.

2.8 Compressive Strength

Compressive tests were carried out on beams with a width and length of 10x20 cm with a height of 10 cm, 8 cm, and 6 cm, respectively. In comparison, the test object is in the form of a cube with a size of 9 cm, 7 cm, and 5 cm. The cube test object is obtained from the beam test object being cut to become a cube per SNI 03-0691-1996 standards. The specimens with variations in shape and height are then tested for compressive strength. This test is carried out to estimate the amount of compressive force that can be resisted by paving blocks before failure. The compressive strength test was tested by curing for 14 and 28 days and then allowed to stand for 24 hours. After that, a compression test using a compression machine with a capacity of 250 kN at a 1 mm/minute speed until it fails. The compressive strength test can be seen in Figure 4. and the following equation is used to calculate the compressive strength:

$$\sigma = \frac{F}{A} \quad (1)$$

where σ is the compressive strength (MPa), F is the compressive load crushed concrete (kN), and A is the surface area (mm²).

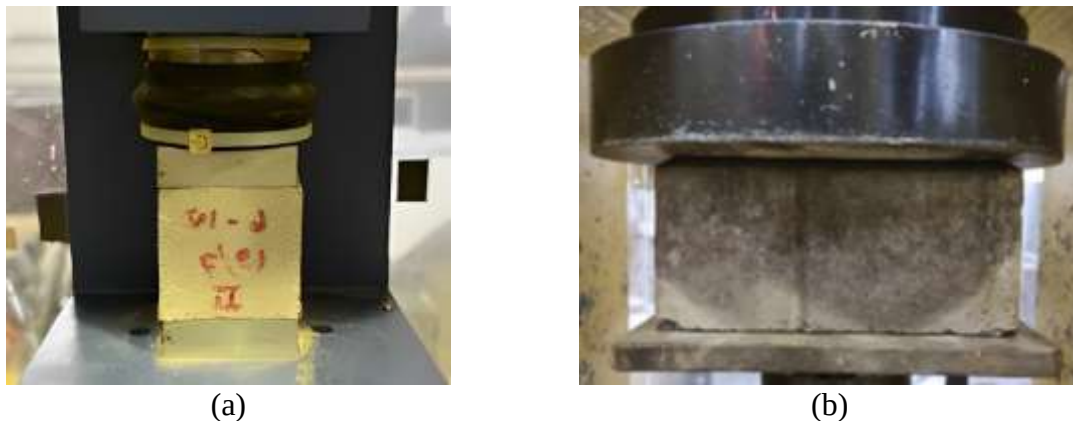


Figure 4: Compressive Strength Test (A) Cube Sample (B) Beam Sample

2.9 Flexural Strength

They tested the flexural strength of asphalt paving blocks using the 3-point method with a test object size of 100 x 100 x 200 mm. Tests were carried out at one-third of the span and the middle of the span at the load above using a compression machine with a load of 0.1 MPa per second. The flexure test was carried out when the concrete paving block was 14 and 28 days old. The following equation is used to calculate flexural strength, and the test scheme can be seen in Figure 5:

$$\text{Flexural Strength, MPa} = \frac{3PL}{2bd^2} \quad (2)$$

where P = load at failure (N); l = span length (mm); b = width of specimen (mm); d = depth of specimen (mm).



Figure 5: Flexural strength test setup

3. Results and Discussion

3.1 Characteristics of The Beam Test and Cube Test

In this study, the aggregate material used RCA to preserve the environment. However, the test results showed that the compressive strength was lower than the compressive strength value of the new aggregate material. This recycled paving block material is then used to compare the compressive strength tests of the beam and cube methods.

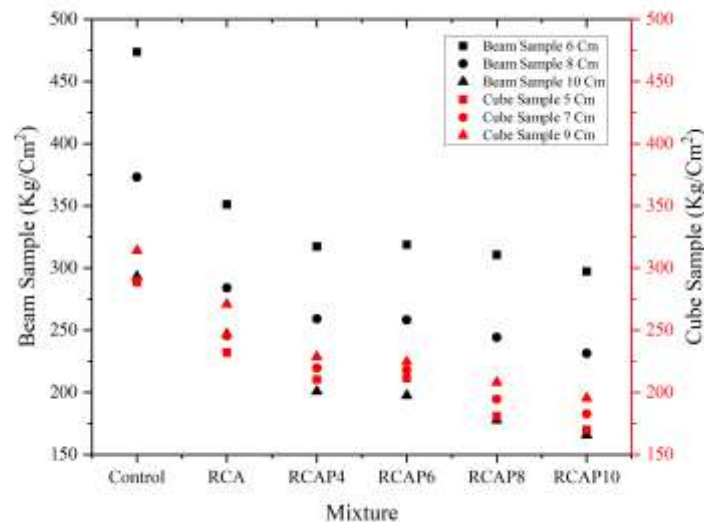


Figure 5: Comparison of Compressive Strength Values of Beam and Cube Samples

These two tests produce different compressive strength characteristics for samples with different thicknesses, namely for sample beams with heights of 6 cm, 8 cm, and 10 cm and sample cubes with heights of 5 cm, 7 cm, and 9 cm. Figure 6 shows the results of a cube-shaped concrete paving block test with the same mixture variation and different heights, showing a strength value where a higher thickness indicates a higher compressive strength value.

On the contrary, the beam test for samples with lower thickness shows higher compressive strength values. Results like this also occur in samples with added plastic material. Whereas cube-shaped concrete paving blocks produce a more realistic value, this is in line with the recommendation of SNI 03-0691-1996 that a cube sample is used to test paving blocks.

Statistical observation of concrete paving block samples in the form of beams cannot be used as a direct reference. Instead, a conversion value is required from the beam test to the cube test, where the amount depends on the thickness of the concrete paving block.

3.2 Effect of Plastic Waste on Paving Block Performance

Flexibility testing was carried out using concrete paving blocks with length, width, and height dimensions of 20 cm, 10 cm, and 10 cm. First, the plastic is treated with a chopping machine, and then the finely chopped pieces are filtered through sieve number 16 with a size of less than 1.16 mm. Mixture variations used are RCA and RCA + Plastic 4%, 6%, and 8%, and Figure 7 shows the flexural strength test results for variations in paving block concrete mixes.

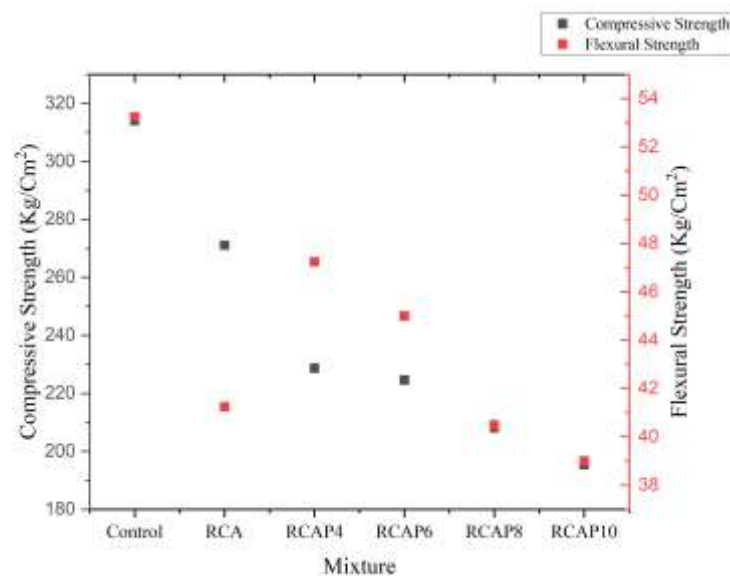


Figure 6: Compressive Strength and Flexural Strength

The test results show a difference in the tendency of the compressive strength and flexural strength test results. The new paving block material (sample control) produces the same value of compressive strength and flexural strength. Whereas in testing the compressive strength of the material with added plastic, it decreases according to the increase in the plastic content in the mixture. Using RCA resulted in lower performance than the sample control test results. This decrease is very dependent on the condition of the RCA material. However, the compressive strength value is still sufficient for road pavement material.

Figure 6 shows that adding 4% and 6% plastic resulted in a higher flexural strength than the plastic RCA + mixture. Using 4% and 6% plastic contributed to better flexural strength and compressive strength characteristics than other RCA compositions. This contribution indicates the addition of 4-5% plastic to using RCA as recycled paving block material. Several studies explain that the research results with the addition of plastic will reduce the compressive strength value (Uvarajan et al. 2022). In other research, it was explained that adding RCA decreased the strength of concrete paving blocks by up to 10% (Luo et al. 2022). Some of the results of these studies align with the research conducted; namely, the addition of RCA resulted in a decrease in strength values.

4. Conclusions

- The results of laboratory tests on compressive strength using beam samples show unrealistic results where thinner beams have higher compressive strength values. In contrast to the compressive strength test using cube samples, where the compressive strength values are more realistic, this follows the recommendation of SNI 03-0691-1996 that cube samples are used for paving block tests.
- RCA substitute paving block concrete produces a lower compressive strength value than new materials. However, 40% of RCA still performs sufficiently as a pavement material for pedestrian facilities or light traffic.

- The use of 4% and 6% plastic waste has shown benefits in increasing the value of flexural strength. This increase indicates that adding plastic can contribute to flexural strength and environmental sustainability.

Acknowledgment

This research was funded by the Direktorat Riset dan Pengabdian Masyarakat Universitas Indonesia through the PUTI Grant 2022, funded by contract no. NKB-714/UN2.RST/HKP.05.00/2022. The authors would like to acknowledge the roles of our bachelor students, Bima A.R and M. Ringga R.H, in assisting during the data collection process of this research.

References

- Arjun Siva Rathan, R. T., V. Sunitha, and V. Anusudha. 2022. "Parametric Study to Investigate the Deflection and Stress Behaviour of Interlocking Concrete Block Pavement." *Road Materials and Pavement Design* 23 (10): 2293–2316. <https://doi.org/10.1080/14680629.2021.1963819>.
- Bajpai, Rishabh, Vedant Soni, Anshuman Shrivastava, and Dipankar Ghosh. 2023. "Experimental Investigation on Paver Blocks of Fly Ash-Based Geopolymer Concrete Containing Silica Fume." *Road Materials and Pavement Design* 24 (1): 138–55. <https://doi.org/10.1080/14680629.2021.2012236>.
- Chu, S. H., Chi Sun Poon, C. S. Lam, and L. Li. 2021. "Effect of Natural and Recycled Aggregate Packing on Properties of Concrete Blocks." *Construction and Building Materials* 278 (April). <https://doi.org/10.1016/j.conbuildmat.2021.122247>.
- Djamaluddin, Abdul Rachman, Muhammad Akbar Caronge, M. W. Tjaronge, Asiyanthi T. Lando, and Rita Irmawaty. 2020. "Evaluation of Sustainable Concrete Paving Blocks Incorporating Processed Waste Tea Ash." *Case Studies in Construction Materials* 12 (June). <https://doi.org/10.1016/j.cscm.2019.e00325>.
- Jamshidi, Ali, Kiyofumi Kurumisawa, Gregory White, Tatsuo Nishizawa, Toshifumi Igarashi, Toyoharu Nawa, and Jize Mao. 2019. "State-of-the-Art of Interlocking Concrete Block Pavement Technology in Japan as a Post-Modern Pavement." *Construction and Building Materials*. Elsevier Ltd. <https://doi.org/10.1016/j.conbuildmat.2018.11.286>.
- Jawaid, Mohammad, Balbir Singh, Lau Kia Kian, Sheikh Ahmad Zaki, and A. M. Radzi. 2023. "Processing Techniques on Plastic Waste Materials for Construction and Building Applications." *Current Opinion in Green and Sustainable Chemistry*. Elsevier B.V. <https://doi.org/10.1016/j.cogsc.2023.100761>.
- Kankam, Charles K, Jonathan Sasah, Kwaku Ansa-Asare, and Bismark K Meisuh. 2018. "Study of Recycled Polyethylene Plastic Waste as Binder in Building Block for Greener Construction Rehabilitation of Distressed Reinforced Concrete Members View Project Blended Cement View Project Bismark K. Meisuh Study of Recycled Polyethylene Plastic Waste as Binder in Building Block for Greener Construction." <https://www.researchgate.net/publication/330262227>.
- Luo, Wenjie, Shu Liu, Yunfeng Hu, Dongdong Hu, Kien Woh Kow, Chengheng Pang, and Bo Li. 2022. "Sustainable Reuse of Excavated Soil and Recycled Concrete Aggregate in Manufacturing Concrete Blocks." *Construction and Building Materials* 342 (August). <https://doi.org/10.1016/j.conbuildmat.2022.127917>.
- Machado da França, Ana Paula, and Fernanda Bianchi Pereira da Costa. 2023. "Evaluating the Effect of Recycled Concrete Aggregate and Sand in Pervious Concrete Paving Blocks." *Road Materials and Pavement Design* 24 (2): 560–77. <https://doi.org/10.1080/14680629.2021.2020680>.

- Makul, Natt, Roman Fediuk, Mugahed Amran, Abdullah M. Zeyad, Sergey Klyuev, Irina Chulkova, Togay Ozbakkaloglu, Nikolai Vatin, Maria Karelina, and Afonso Azevedo. 2021. "Design Strategy for Recycled Aggregate Concrete: A Review of Status and Future Perspectives." *Crystals*. MDPI AG. <https://doi.org/10.3390/cryst11060695>.
- Mampearachchi, Wasantha. 2019. *Handbook on Concrete Block Paving*. *Handbook on Concrete Block Paving*. Springer Singapore. <https://doi.org/10.1007/978-981-13-8417-2>.
- Meesaraganda, L. V.Prasad, and V. S.Prasad Kakumani. 2021. "Effect of Various Combinations of Aperture Diameter and Pattern on Concrete Paver Block." In *Materials Today: Proceedings*, 45:5494–99. Elsevier Ltd. <https://doi.org/10.1016/j.matpr.2021.02.201>.
- Salehi, Safoura, Mehrdad Arashpour, Jayantha Kodikara, and Ross Guppy. 2021. "Sustainable Pavement Construction: A Systematic Literature Review of Environmental and Economic Analysis of Recycled Materials." *Journal of Cleaner Production*. Elsevier Ltd. <https://doi.org/10.1016/j.jclepro.2021.127936>.
- Uvarajan, Turkeswari, Paran Gani, Ng Chuck Chuan, and Nur Hanis Zulkernain. 2022. "Reusing Plastic Waste in the Production of Bricks and Paving Blocks: A Review." *European Journal of Environmental and Civil Engineering*. Taylor and Francis Ltd. <https://doi.org/10.1080/19648189.2021.1967201>.
- Yeo, Jerome Song, Suhana Koting, Chiu Chuen Onn, and Kim Hung Mo. n.d. "An Overview on the Properties of Eco-Friendly Concrete Paving Blocks Incorporating Selected Waste Materials as Aggregate." <https://doi.org/10.1007/s11356-021-13836-3>/Published.
- Yeo, Jerome Song, Suhana Koting, Chiu Chuen Onn, Mohammed K.H. Radwan, Chee Ban Cheah, and Kim Hung Mo. 2023. "Optimisation and Environmental Impact Analysis of Green Dry Mix Mortar Paving Block Incorporating High Volume Recycled Waste Glass and Ground Granulated Blast Furnace Slag." *Environmental Science and Pollution Research*. <https://doi.org/10.1007/s11356-023-26496-2>.

ELEMENTAL ANALYSIS OF PRINTING INK WITH THE APPLICATION OF LASER-INDUCED BREAKDOWN SPECTROSCOPY (LIBS)- A REVIEW

Adlina Syafura Ahmad Sabri¹

Hamizah Md Rasid¹

Reena Abd Rashid¹

Umi Kalsum Abdul Karim²

Mohamed Sazif Mohamed Subri³

Mohamed Izzharif Abdul Halim^{1*}

¹ School of Chemistry & Environment, Faculty of Applied Sciences, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia

² Faculty of Applied Sciences, Universiti Teknologi MARA, 35400 Tapah Road, Perak, Malaysia

³ Forensic Laboratory, Criminal Investigation Department, Royal Malaysian Police, 50560 Bukit Aman, Kuala Lumpur, Malaysia

Abstract: Forgery cases involving printed documents have recently increased due to advanced printing technologies. As a result, various analytical instruments are currently applied in forensic document examination to analyse printing ink. Analyses of printing ink are primarily focused on the ink's chemical components, namely colourants, vehicles, and additives. Previous studies on ink analysis show an increasing trend in the use of spectroscopy methods including the application of Laser-Induced Breakdown Spectroscopy (LIBS). This instrument is favoured by document examiners in analysing printing ink due to its versatility, simple preparation, multi-elemental detection, minimal destruction, and precise readings. The examination of printing ink using LIBS is reviewed in this publication in addition to findings from previous research. The strengths and drawbacks of this technique, as well as contemporary approaches, are also explored in this article.

Keywords: Document Examination, Ink Analysis, Printing Ink, Spectroscopy Technique, Laser-Induced Breakdown Spectroscopy.

1. Introduction

The growing market and demand for advanced printers, computers, scanners, and photocopiers have fuelled the progress of printing technology. This progress has increased the forgeries of documents in printed form, such as academic certificates, currency, medical records, insurance papers, wills, contracts as well as travel and identity documents [1-8]. Therefore, ink analysis of printed documents has become indispensable in verifying questioned documents. Questioned documents refer to any form of written or printed texts on various substrates (paper, banknotes, etc.) with questionable authenticity and these documents can be treated as evidence in civil or criminal cases [9]. The first analysis of printing ink in 1914 was performed to determine ink composition and detect harmful materials [10]. High-speed computer printers were invented in 1953 causing a large influx of printers for personal or commercial uses [11-16]. As a result, the trend of ink analysis from printing inks took off in the 1960s [11-16].

Generally, printed documents can be categorised based on their ink type namely toner, inkjet, offset, and intaglio. Printing ink contains vehicles (resin and/or solvent) with flowing or drying characteristics that dissolve or suspend the colourants (pigments or dyes) [8,17-19]. Inorganic and organic additives are added to the mixture contributing specific properties, modifying the drying process, influencing the ink's actual appearance, or controlling the density or flowing characteristic of the ink [19,20]. As a result, a wide variety of substances, according to the ink's type, brand, and batch, provide unique chemical signatures which are beneficial to forensic scientists [3,21]. Although the exact chemical composition of a particular ink is undisclosed by its manufacturers, each ink variant has a unique fingerprint in terms of its analytical makeup.

A Forensic Document Examiner (FDE) plays a role to ascertain the authenticity of documents by detecting alterations, deletions, and obliterations as well as establishing the exact point in the document's printing time. Thus, FDE is encouraged to apply analytical techniques to provide a highly-accurate ink analysis. Previous studies on printing ink depended on chemical or elemental compositions of ink, intersecting printing lines, ink dating, paper analysis, etc. FDE commonly employs visual examinations such as microscopic analysis and light examination with Visual Spectral Comparator (VSC) [6,22,23]. However, these techniques are subjected to human error and the results are inadequate to investigate the authenticity of the documents [22,23]. Hence, it is necessary to apply further chemical analysis for improved characterisation and discrimination of the ink.

Numerous analytical approaches (destructive and non-destructive) have their purposes in printing ink analysis as reported in the scientific literature [20]. Destructive techniques are time-consuming with delicate sample preparation, which eventually damages the samples. Despite their destructive nature, these techniques are still practised in forensic analysis due to their powerful compound identification in ink [3,14,21-28]. Some examples of destructive techniques are mass spectrometry, capillary electrophoresis, and High-Performance Thin Layer Chromatography (HPTLC). In contrast, non-destructive techniques require minimal sample preparation with the added advantages of multiple measurements and preservation of the documents' integrity [4]. These techniques involve utilising spectroscopy instruments that are occasionally used in tandem and in conjunction with multivariate data analysis. Infrared (IR), Raman, and elemental spectroscopy are a few examples of non-destructive techniques [4,25,29-37].

Laser-Induced Breakdown Spectroscopy (LIBS) is an atomic emission spectroscopy that applies the use of laser ablation. LIBS is first applied by Brech and Crossin [38] and along the way, LIBS has found its way into other fields of study, such as biology [39-43], geology [44], astronomy [45], archaeology [46-49], and artistic prints [47,52-54]. LIBS is a semi-destructive elemental analysis that employs a surface technique using laser ablation with high sample throughput [8,55-57]. This cost-effective instrument collects information from the ablated mass with high sensitivity and precision [6,8,55,56]. To date, only seven studies on the application of LIBS in printing ink analysis exist due to the infancy of LIBS in this area of forensic interest. Therefore, this paper discusses the pros and cons of LIBS in printing ink analysis by reviewing the results from previous research in order to determine the aptness of this technique for printing ink. The hybrid of LIBS with other instruments and its role in multivariate data analysis are also discussed.

2. LIBS' Principles

LIBS is considered a relatively recent innovation in forensic science thus, its features such as lasers, spectrometers, and detectors are constantly being upgraded [55]. LIBS is compartmentalised into several subsystems; a pulsed laser focused by lenses on the target, a delay pulse generator, an ablation cell for the plasma formation, an optical system to collect emitted signal, a spectrometer with a sensitive detector, and a computer [8,57-60]. Besides the type of laser, the configurations of the spectrometer-detector and time control unit, environmental factors should be considered to improve any experiment with LIBS [61]. Figure 1 shows a typical setup for LIBS.

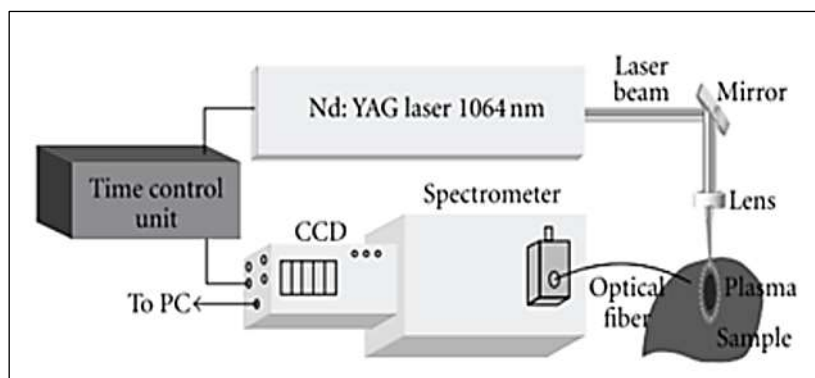


Figure 40: Typical Setup of LIBS [61]

A highly energetic laser-pulsed beam (in nanosecond ranges) is aimed at the sample surface to provoke optical sample excitation and induce the plasma [3,58,61,62]. Consequently, this light energy causes the ablation of a small sample mass and the creation of high-temperature plasma. The temperature of the sample surface rapidly increases up to the vapourisation threshold [59,61]. This process causes phase change at the sample surface and several other phenomena, such as photoelectron emission and plasma formation [59]. The parameters of laser-produced plasma are determined by the excitation temperature and electron density [63]. The plasma with excited atoms and ions of the analysed sample produces discrete lines, bands, and an overlying continuum [3,59,60,61]. The discrete lines consist of intensity, shape, and wavelength which help characterise the samples [3,61]. The laser energy directly affects plasma emission spectra and the signal-to-noise ratio [64].

Besides laser energy, the laser wavelength is crucial in the formation of plasma [63]. The suitable wavelength is selected depending on the expected analytical work with LIBS [63]. Besides that, the time delay parameter in LIBS is as important as the laser energy due to the dynamic processes of plasma formation and radiation [64]. Time delay allows the decay of the continuum radiation before the spectral measurements [60,64]. The spectra are affected by time variation in which the dissipation of the laser plasma and the formation of the spectrum occurs in microseconds [60]. The process of selecting time delay (gate delay) and integration time (gate window) must be based on the laser wavelength, which subsequently affects the signal-to-background ratio [61]. An optimised delay can minimise interferences and improve signal intensity however, a longer time delay can cause lower signal intensities [56,64].

The plasma emission is directed to a spectrometer (e.g., Czerny-Turner and Echelle) with optical fibres and/or lenses [58,61]. The spectrometer is used to diffract the collected light, which then is dispersed by a grating, thereby producing “spectral signatures” of the sample [8,61,62]. Subsequently, a sensitive detector detects the light, such as a photomultiplier tube

(PMT), a photodiode array (PDA), or a charge-coupled device (CCD) [8,61]. Finally, the computer connected to the system is used for data storage and processing of the acquired spectrum. The raw data (intensity versus wavelength) can be obtained from the spectrum with spectral lines that are formed based on elements present in the samples [55]. The spectral lines are typically shown in wavelengths ranging from 190 to 900 nm. The National Institute of Standards and Technology (NIST) database can be used for comparison of the wavelengths [8,16,55,65]. Additionally, simultaneous quantification is possible and the concentration of the elements can be determined based on their emission intensity [8].

2.1 Stages of Plasma Lifetime in LIBS

Optimising the parameters for LIBS measurements requires an understanding of LIBS' plasma physics [61]. Consideration of the environmental factors that may affect the plasma lifetime and features is also necessary [61]. The plasma lifetime is divided into three stages, as seen in Figure 2 below [61]. Several processes occur as the laser evokes the creation of plasma. It begins with the emission of continuous Bremsstrahlung [59,63]. Plasma ignition is the first stage of plasma lifetime occurring during the laser pulse which breaks the bond and shields the plasma. The interaction of laser and matter causes the absorption of photons by atoms. This absorption causes the electrons to acquire higher energy levels [61]. This process is affected by the duration of the laser pulse, irradiance, and laser type. The plasma becomes opaque for the last leg of laser radiation during the plasma shielding effect. At this stage, the laser pulse is either reflected or absorbed by the plasma surface. The radiation does not reach the sample surface which prevents the ionisation of the material and reduces the ablation rate. Then, the craters are formed with melted and deposited material around them. Simultaneously, the size and lifetime of the plasma are multiplied when the plasma is reheated.

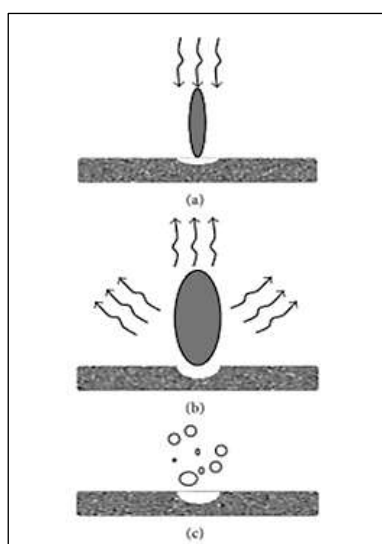


Figure 41: Stages of Plasma Lifetime Which Are (a) Ignition of Plasma, (b) Expansion and Cooling of Plasma, and (c) Ejection of Particle and Condensation [61]

The following stage is the expansion and cooling of the plasma. The factors that influence this stage are spot size, environmental aspects, ablated mass, and energy coupled to the plasma [61]. Atomic emission occurs during the cooling process thus, this stage is vital to optimising the acquisition of LIBS spectra. Moreover, recombining electrons and ions in this stage is compulsory to give rise to line emission, which influences the radiation processes [63]. After

the ignition process, plasma expansion, and cooling continue meanwhile, electron temperature and density change simultaneously. As the electrons downshift to lower energy levels, the electrons emit photons in the decay. This process is synonymous with the de-excitation of atoms [61].

Line emissions are more visible after plasma formation as the plasma cools down. However, line emissions are hidden in the continuous emission due to radiative recombination and the Bremsstrahlung effect [59,61]. The plasma radiation first appears as a continuous spectrum that fades quickly in the wake of plasma relaxation [59,63]. This radiation constitutes atomic lines from decays of excited atoms, ions, and small molecules [59,63]. The distinct energy levels of different atoms cause them to emit characteristic photon energies with narrowband emissions forming the spectral emission lines in LIBS spectra. Therefore, a time delay in the time control unit is required to avoid the continuum radiation [59,61]. The final stage of the plasma lifetime is particle ejection and condensation which are insignificant for LIBS measurements [61]. A portion of the materials is ablated as particles that cause the exfoliation of a solid sample, formation of condensed vapour, and ejection of a liquid sample.

3. Instrumental Setup of LIBS in Previous Studies

The setup of LIBS involves selecting the best laser source as the light energy and the configuration of laser parameters such as wavelength, pulse time, and laser energy. Optimisation and calibration of LIBS parameters are imperative to acquire the highest intensity of signals without saturating the spectrum [2,56]. Due to the resultant craters after laser ablation, the size of the craters should be considered to ensure the least amount of ablated substrate on which the ink was printed. Previous research's instrumental setup and parameters of LIBS for printing ink analysis are summarised in Table 1.

Table 1: List of Instrumental Setup of LIBS in Previous Studies

Type of sample	Laser source (wavelength; energy; type)	Spectrometer	Time delay	Detector	Spot size	Additional notes	Ref.
Ink on paper	1064 nm; 170 mJ; Nd:YAG	Ocean Optics LIBS2500+ spectrometer	2.6 μ s	Not available	300 μ m	The sample chamber was purged with argon gas for 60 s for air removal and increased spectral intensity. The analysis mode was line mode.	[1]
Ink on paper	1064 nm; 270 mJ (10 ns); Nd:YAG	Ocean Optics HR4000 spectrometer	Not available	Not available	Not available	The research was performed in ambient air using 60 μ m UV-Vis optic fibre. Integration time was fixed at 100	[2]

ms for a laser
voltage of 900 mJ.

Ink on paper	1064 nm; Not available; Nd:YAG	6-channel broadband spectrograph (190 – 1040 nm)	1.4 μ s (inkjet) & 1.2 μ s (toner)	CCD linear array	300 μ m	35% energy was used for 150 shots with a 25 μ m/s speed rate at a 2 Hz frequency and 1.1 ms of integration time.	[7]
Ink on banknote papers	1064 nm (spectra collected at 185 – 904 nm UV-Vis ranges); 150 mJ; Nd:YAG	6-channel Czerny-Turner	Not available	CCD	0.6 to 1 mm	The experiment was performed at atmospheric pressure by running the laser for 6s. The spectral resolution was around 0.1 nm.	[56]
Ink on pharmaceutical packages	266 nm; 5.4 mJ; Nd:YAG	6-channel Czerny-Turner (190 – 1040 nm ranges)	0.9 μ s	Not available	100 μ m	The ablation mode used a straight-line mode.	[62]
Ink on paper	198 – 318 & 345 – 888 nm (UV & Vis ranges, respectively); 10 mJ; Nd:GGG	Two-channel fibre optic spectrometer	40 μ s (double-pulse mode)	CCD	Not available	A micro-LIBS system was used with two consecutive laser pulses that have similar energies with an integration time fixed at 2 ms. The measurements were performed under argon gas.	[66]
Ink on paper	1064 nm; 4.1 mJ (offset & intaglio inks), 13 mJ (toner), 7.8 mJ (inkjet); Nd:YAG	6-channel broadband spectrometer	0.1 μ s (inkjet, offset & intaglio) & 0.8 μ s (toner)	CCD linear array	200 μ m	Laser Ablation – Laser Inductively Coupled Plasma – Mass Spectrometry (LA-ICP-MS) was used in tandem with LIBS. The ablation mode	[67]

was single-line
mode.

As seen in Table 1, the most common laser used in previous studies was Nd:YAG [1,2,7,56,62,67]. This laser type with a fundamental wavelength of 1064 nm is the preferred light source in LIBS as it can quickly generate compact plasma [61]. As highlighted by Anabitarte *et al.* [61], a higher laser wavelength produces a lower ablation rate with a higher elemental fractionation. A study by Metzinger *et al.* [66] used a different laser type (Nd:GGG) due to a change in the LIBS system. The intensities of LIBS signal emission were increased by configuring the emission of laser pulses (Table 1). Several authors used Czerny-Turner [56,62] and Ocean Optics [1,2] spectrometers in their studies. Czerny-Turner spectrometer is common in LIBS which consists of a diffraction grating, an entrance slit, and two mirrors [61]. The slit is passed through with the light which hits the first mirror. The mirror pivots the light onto the diffraction grating before the light is directed onto the focal plane with the detector by the second mirror.

The application of the instrument determines the suitable detectors to be used. CCD was equipped by several authors as the detector for printing ink analysis [7,56,66,67], while the rest did not divulge details on the type of detector [1,2,62]. CCD detectors provide less background signal for two-dimensional spatial information [61]. As for ablation spot size, several authors measured its size, and they observed spot size on paper samples (200 – 300 μm) was smaller compared to banknote papers (0.6 – 1.0 mm) [1,7,67]. Subedi *et al.* [67] explored the use of a tandem system of LIBS with LA-ICP-MS. As explained by the authors, sample excitation by the emitted light is collected by a fibre optic in the LIBS plasma to a CCD spectrometer. The ablated particles were transported by an argon carrier to the inductively coupled plasma (ICP). The laser energy was varied by the authors according to the types of ink in which inkjet ink consumed the highest energy (13 mJ). Additionally, laser frequency was varied according to ink type. A higher frequency was used for toner ink (1.8 Hz), offset ink (4 Hz), and intaglio ink (4 Hz) compared to inkjet ink (0.8 Hz) [67]. High frequency can cause more damage to the samples. Therefore, low frequency is more suitable for inkjet ink to minimise interferences from paper since this ink type is immersed in the paper.

4. Elemental Analysis of Printing Ink

LIBS is a compact and stand-alone system that is capable of simultaneous rapid multi-elemental detection and surface analysis (<1 second per analysis) [7,56-58,60,68]. This instrument can analyse any material in various physical states i.e., solids, liquids, gas etc. [60,64,68]. There is no requirement to prepare or treat the sample before analysis. This non-existent prerequisite preserves the sample integrity, which is favourable for ink samples [3,60]. LIBS can perform atomic-level quantitative and qualitative chemical analyses by properly identifying a specific element (in an ionised or neutral state) for each emission line [61]. Moreover, pre-set elemental menu is not necessary for LIBS, unlike in LA-ICP-MS. Broadband scanning can be performed using LIBS and it is useful for the detection of unusual peaks [7,67]. In printing ink analysis, LIBS can provide inorganic information about the ink. This feature is a rarity in standard techniques [66]. Qualitative analysis can be carried out accurately as it detects the atomic emission of all elements in the analysed sample [61].

Quantitative analysis of LIBS determines the relative number of elements in units of ppm (concentration of elements), ng (absolute mass of elements), or ng/cm² (surface concentration) [61]. This analysis encompasses the non-calibration method and calibration method. The latter uses a calibration curve tied to experimental factors [61,69]. However, the optimisation and development of LIBS are time-consuming [1]. Furthermore, LIBS may not be sensitive enough in the quantitative analysis compared to other atomic emission spectroscopies [60]. Additionally, spectral noise can interfere with elemental analysis which can be caused by the presence or absence of distinct emission lines [56]. Nonetheless, generating reproducible spectra that show consistent qualitative differences through a selection of valid spectral lines can provide a solution to this problem [56].

Another difficulty of using LIBS is interpreting multivariable data in large amounts within a short period of time [2,55]. A reliable overall data analysis could not be achieved when interpreting an abundance of spectral lines. However, multivariate data analysis can be applied for a systematic classification and differentiation of the samples. LIBS is also a semi-destructive instrument, leaving a small visible crater after the laser ablation [2,3,6,57]. Metzinger *et al.* [66] microscopically scrutinised the images of before and after LIBS measurements on inkjet and toner ink on paper. They observed that the diameter of all craters after ablation was tenfold or hundredfold bigger in size compared to the size of all prints which was only a few μm in diameter. Deep craters were observed on inkjet ink with 30 μm to 95 μm in diameter. The low density of the fibrous structure of the paper causes higher penetration of laser on inkjet ink. For toner ink, the ablation only affected the toner layer since the light energy is highly absorbed in the layer. The toner particles had a weak adhesion to the paper fibre, and their heat conduction was higher than the materials in the paper. Despite these disadvantages, several authors argued that LIBS is a viable technique to distinguish printing ink from various sources [2,7].

4.1 Detected Elements in Previous Studies

The elemental composition of printing ink can vary due to the mixture of chemical components created during the manufacturing process [7]. Instead of a complete quantitative method, qualitative and semi-quantitative data from intensities of the peak are favoured for samples of toner and inkjet ink. The detected elements in previous studies are listed in Table 2. Most studies provide detailed information on types of ink, except the research conducted by Hui *et al.* [2]. The most analysed ink types are toner and inkjet ink printed on paper as substrate [1,2,7,66,67]. Only two studies analysed printing ink on banknote papers [56] and pharmaceutical packages [62].

Table 2: List of Elements Found in Printing Ink

Sample	Elements														Ref.
	Al	Ba	Ca	Cd	Co	Cr	Cu	Fe	Hf	I	K	Mg	Mn	Na	
Toner and inkjet ink on paper	×	×	×	×	×	×	×	/	×	×	×	×	×	×	[2]
Toner ink on paper	×	×	/	×	×	/	/	×	/	×	×	/	×	×	[1]

Toner ink on paper	/	×	×	×	/	×	/	/	×	/	/	/	/	/	[7]
Toner ink on paper	×	×	/	×	×	×	×	/	×	×	×	×	×	/	[66]
Toner ink on paper	×	×	/	×	×	×	×	/	×	×	/	×	/	/	[67]
Inkjet ink on paper	×	×	/	×	×	/	/	×	×	×	×	/	×	×	[1]
Inkjet ink on paper	×	×	×	×	/	×	/	×	×	×	/	/	×	×	[7]
Inkjet ink on paper	/	×	/	×	×	×	×	×	×	×	/	×	×	/	[67]
Offset ink on paper	/	/	/	×	/	×	/	×	×	×	/	/	/	/	[67]
Intaglio ink on paper	/	×	/	×	/	×	/	×	×	×	/	/	/	/	[67]
Ink on banknote papers	/	×	/	×	/	/	×	/	×	×	/	×	/	/	[56]
Ink on pharmaceutical packages	/	/	/	/	×	/	/	/	×	×	/	/	/	×	[62]

/ denotes the presence of the element in the sample

× denotes the absence of the element in the sample

Table 2: List of Elements Found in Printing Ink (Continued)

Sample	Elements															Ref.
	Nb	Ni	Pd	Rb	Rh	S	Sb	Sc	Si	Sr	Ti	V	Y	Zn	Zr	
Toner and inkjet ink on paper	×	×	×	×	×	×	×	×	×	/	/	×	×	×	×	[2]
Toner ink on paper	×	×	/	×	/	×	×	×	/	/	/	/	×	×	×	[1]
Toner ink on paper	×	/	×	/	×	/	/	/	/	/	/	/	/	/	/	[7]
Toner ink on paper	×	×	×	×	×	×	×	×	×	×	/	×	×	×	×	[66]

Toner ink on paper	×	×	×	×	×	×	×	×	/	×	×	×	×	×	×	[67]
Inkjet ink on paper	×	×	×	×	×	×	×	×	/	/	×	×	×	×	×	[1]
Inkjet ink on paper	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	[7]
Inkjet ink on paper	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	[67]
Offset ink on paper	×	×	×	×	×	×	×	×	/	/	/	×	×	×	×	[67]
Intaglio ink on paper	/	×	×	×	×	×	×	×	×	/	/	×	×	×	/	[67]
Ink on banknote	×	/	×	×	×	×	×	×	×	×	/	×	×	×	/	[56]
Ink on pharmaceutical packages	×	/	×	×	×	×	×	×	/	/	/	×	×	×	×	[62]

/ denotes the presence of the element in the sample

× denotes the absence of the element in the sample

Several authors employed the conventional technique of spectral overlay as the data acquisition approach [2,7,62,67]. According to Trejos *et al.* [7], errors can be minimised using this approach when spectral regions with below detection limit and linear response are used for comparison. Two independent examiners reviewed the spectral data for similarities and differences of the replicated spectra to differentiate the samples [7,25]. Acquired spectra were first divided into smaller regions of interest. The signals between replicates for each region were normalised with background subtraction in LIBS software [2,7]. Elements beyond the detection limit were identified with peak search tolerance comparable to spectra resolution. This step marked the variability within the samples by comparing replicates of the samples. At least two spectral lines for each element were selected for visual comparison and confirmation [2,7,25]. The researchers then observed significant differences between the samples, such as spectral shapes and their relative peak heights. The samples were discriminated if any significant difference was detected [2,7].

As shown in Table 2, inkjet ink samples contain fewer elements compared to other types of ink [19,56,66]. Several authors identified similar elements in inkjet ink namely Ca, Mg, Cu, and K [1,7,67]. The element of Li was observed to be present in inkjet ink but this element was absent in other types of ink. Li plays a vital role in inkjet printing for electrical conductivity [7,67]. In

addition, elements namely Al, Co, Sr, Si, Cr, and Na were discovered by several authors in inkjet ink [1,7,67]. Chemical formulations used in inkjet ink may cause fewer elements to be detected in this ink than in toner ink [67]. Furthermore, inkjet ink was partially absorbed in paper fibre. This absorption caused poor discrimination due to interferences from high paper contribution [66].

In contrast, toner ink was deposited onto the paper surface. Therefore, this ink type was discriminated efficiently [66]. A tandem LIBS-LA-ICP-MS discovered a similar observation in which inkjet ink was primarily embedded in the paper fibre while the layers of other types of ink (intaglio, offset, and toner ink) were slightly raised above the paper [67]. Lennard *et al.* [1] selected emission lines for detected elements in toner ink samples namely Ca I (643.91 nm), Cr I (520.84 nm), Cu I (327.36 nm), Hf II (325.37 nm), Mg I (285.21 nm), Pd I (324.27 nm), Rh I (332.3 nm), Si I (288.21 nm), Sr II (460.73 nm), Ti I (498.17 nm), and V I (440.82 nm). As for Hui *et al.* [2], only several spectra regions of interest (ranging from 220 nm to 500 nm) were used to detect emission lines from ink compositions to reduce the contribution from paper spectra. Figure 3 shows the peaks of Fe II and Mg II appearing in regions of 220 nm to 280 nm [2]. A minimal paper contribution was observed for Fe II, but the paper contribution was similar to the ink component for Mg II [2].

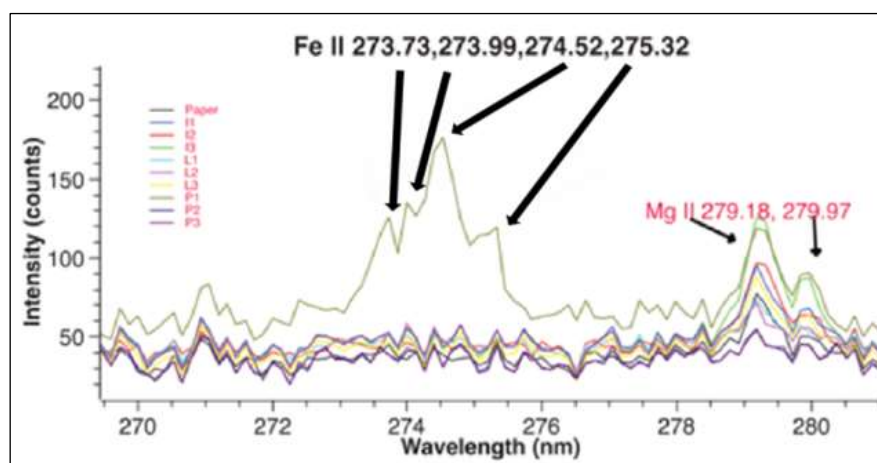


Figure 42: Emission Lines of Elements Fe II and Mg II in Samples with A Paper Contribution

Moreover, only quantitative differences were observed for all toner ink samples except for one photocopier sample that showed qualitative differences [2]. This sample showed unique lines of Sr at wavelengths 407.41 nm, 460.73 nm, and 421.27 nm. Trejos *et al.* [7] detected 20 elements in toner ink which was the highest number of elements compared to other studies. The authors detected I, Ni, Rb, Sb, Sc, S, V, and Y in toner ink samples that went undetected in other studies [7]. Meanwhile, only Lennard *et al.* [1] detected Rh, Pd, and Hf. As reported by Metzinger *et al.* [66], carbon-based particles or soot were commonly found in most black toner ink samples, but iron oxide was present in one of the samples. Fewer lines and lower intensity in the structured background were observed in UV spectra compared to Vis spectra as explained by the authors. Furthermore, the detection of Zr, Nb, and Hf in toner ink samples is expected since these elements had been included in several patents and books for toner ink formulations [70,71]. Subedi *et al.* [67] analysed offset and intaglio inks. They postulated that

each of these ink types had a single element that was absent in other ink. As for offset ink, Ba was detected where else Nb was discovered in intaglio inks only [67].

Król *et al.* [56] proved the presence of Mn as the most frequent element in the banknote papers with the addition of other elements namely Al, Ca, Na, and Ti. Banknote papers contain the element of Ti due to the addition of titanium oxide during manufacturing. Authentic and counterfeit banknotes have different ratios of elemental profiles. Furthermore, the spectra of banknotes can be distinguished from standard office paper based on the number of lines in the spectra. The spectrum of banknotes showed more lines arising from singly charged ions, molecules, and neutral species compared to the office paper spectrum. Previous studies have proven that counterfeit banknotes are printed on standard office paper [56]. The spectral differences between banknotes and office paper are the basis to tell counterfeit banknotes apart from authentic banknotes. In addition, the authors studied the discrepancies between new banknotes and used banknotes. Qualitative differences can be observed in the spectra of both banknotes in several samples. Several used banknotes showed additional emission lines from Fe I at wavelengths from 371.965 nm to 374.942 nm, as shown in Figure 4. These peaks may arise from contamination due to contact with human skin and metallic parts of items for instance, pockets, wallets, and purses.

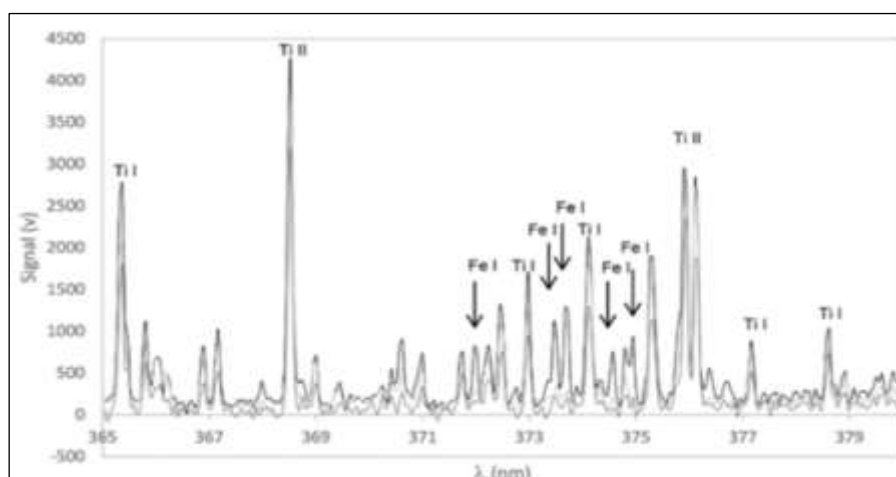


Figure 43: LIBS Spectrum of Used Banknotes with Additional Lines of Fe I As Contamination [56]

Analysis of printing ink on pharmaceutical packages by Haase *et al.* [62] discovered qualitative and quantitative differences between authentic and counterfeit packages. The authors described quantitative differences in signal intensities for peaks of Ca at 397 nm and K at 766 nm for P1 sample (authentic package) and C4 sample (counterfeit package). Samples P1 and C4 also showed qualitative differences with peaks of Mg (285 nm), Si (288 nm), and Al (396 nm). These elements were only found in the authentic package. As seen in Table-2, most of the detected elements (Al, Ca, Cr, Cu etc.) were similar to ink on paper except for Cd, which was detected on pharmaceutical packages only.

5. Comparison of LIBS with Other Spectroscopic Techniques

Several studies applied LIBS in conjunction with other spectroscopic instruments such as Attenuated Total Reflectance – Fourier Transform (ATR-FTIR) spectroscopy, Micro X-ray

Fluorescence (μ XRF) analysis, and LA-ICP-MS. Table 3 summarises the comparison of LIBS with these techniques for the analysis of printing ink.

Table 3: Comparison of LIBS with Other Spectroscopic Techniques in Analysing Printing Ink

Other spectroscopic technique	Sample	Findings	Ref.
LA-ICP-MS	Toner, inkjet, offset, and intaglio ink on paper	LIBS could not detect all elements found by LA-ICP-MS. However, LIBS detected problematic elements (i.e., Si, Ca, K) and additional elements that can discriminate indistinguishable samples in LA-ICP-MS. Lesser interferences and improved signal-to-noise ratio for Fe, Ca, K, and Si in LIBS. Both techniques showed comparable precision figures for inkjet ink (10-15%) but poorer results for toner ink (3-27%). Tandem LIBS-LA-ICP-MS system reduced interferences and improved discrimination ability.	[7,67]
μ XRF analysis	Ink on banknote papers	The same measuring area showed strongly varied elemental compositions from μ XRF analysis and LIBS.	[56]
ATR-FTIR spectroscopy	Ink on pharmaceutical packages	Information from LIBS and ATR-FTIR were complementary. LIBS provided inorganic elemental information. ATR-FTIR revealed the presence of organic compounds.	[62]

Subedi *et al.* [67] built a tandem LIBS-LA-ICP-MS for the analysis of toner, inkjet, intaglio, and offset ink (Table 3). This hybrid enhanced discrimination ability and diminished interferences. The results from this tandem system presented a thorough chemical composition of printing ink. The authors further explained that the optimisation was performed differently in standalone mode compared to tandem mode due to different principles of the techniques. In LIBS, the excitation of atoms and ions and the correct timing for emission line acquisition are crucial for optimisation. As for LA-ICP-MS, the critical factors are sub-micron-sized particles' production and efficient transport. The authors prioritised LIBS optimisation before adjusting the parameters to LA-ICP-MS for tandem mode. Consequently, some sensitivity of ICP-MS was withdrawn to encourage the improved performance of both methods.

Elemental information provided by tandem LIBS-LA-ICP-MS was complementary, which could not be obtained when these instruments were used individually. Additional elements were detected by LIBS that discriminated indistinguishable samples in LA-ICP-MS and LIBS observed fewer interferences with better signal-to-noise and reproducibility for Fe, Ca, K, and Si. Moreover, LIBS successfully detected certain elements (Si, Ca, and K) that are problematic as a result of isobaric interferences in LA-ICP-MS [7,67]. LIBS could not detect all elements due to its lower sensitivity, as opposed to LA-ICP-MS. This fact underscores the superiority of LA-ICP-MS over LIBS. Nevertheless, LIBS is an alternative method for analysing printing ink due to its rapid testing and low associated costs [7]. In terms of precision, both techniques

provided comparable precision figures for inkjet ink (10% - 15%) but toner ink (3% - 27%) gave poorer results due to its larger variations.

Despite low repeatability between measurements, good discrimination for these inks can be accomplished with both techniques. Haase *et al.* [62] applied ATR-FTIR spectroscopy and LIBS to analyse pharmaceutical packages. They discovered that the information from ATR-FTIR spectroscopy was complementary to LIBS. ATR-FTIR provided information from organic constituents, while LIBS revealed the presence of inorganic elements in the samples. However, several peaks from the inorganic composition can appear in ATR-FTIR spectra in the regions from 1000 cm^{-1} to 400 cm^{-1} . Król *et al.* [56] compared the performance of LIBS against μXRF in the ink analysis of banknote papers. They determined that the same measuring area of banknote papers displayed strongly varied elemental compositions from each instrument. The distinction between these two techniques was likely caused by the differences in penetration depth and sensitivity.

6. Merging LIBS with Multivariate Data Analysis

Multivariate data analysis is the application of statistical tools combined with scientific studies to improve sample discrimination and produce extensive data. Manual processing of datasets from LIBS can be impractical due to the combined complexities and sheer volume of chemical information in the spectra (sometimes more than 500 spectra lines), especially with many samples [55,58,62,69,72]. However, systematic classification of samples can be accomplished by introducing chemometric analysis to the data [62,69]. The chemometric analysis can extract valuable information based on the differences among the acquired spectra. This analysis is advantageous for printing ink analysis since it reveals minimal differences between ink samples with the return of higher accuracy for ink discrimination. A summary of applied statistical analyses in previous studies is shown in Table 4.

Table 4: List of Statistical Analyses Applied with LIBS For Printing Ink Analysis

Statistical analysis	Findings	Ref.
Principal Component Analysis (PCA)	Good separation and variation (beyond 95%) were observed for toner and inkjet ink. Significant information from ink manufacturers and a large sampling collection were necessary for sample classification. PCA reduced variables prior to classifications with supervised machine learning models.	[1,2,7,62]
3-sigma criterion	Excellent discrimination was achieved for toner ink samples (97%) and inkjet ink samples (93%).	[1]
Pairwise comparison analysis	Excellent discrimination for toner (98%) and inkjet (100%) samples was achieved using Analysis of Variance (ANOVA) with Tukey's post hoc test.	[1]
t-test	Three indistinguishable pairs of toner samples were discriminated with a 95% confidence limit.	[1]
Classification models	Robust classification models of Linear Discriminant Analysis (LDA) and <i>k</i> -Nearest Neighbor (KNN) were obtained using data fusion of ATR-FTIR and LIBS with low confusion rates and error. The models showed high accuracy with correct classification rates for LDA (85-99%) and KNN (90-100%).	[62]

Discriminant analysis	Advanced multivariate discriminant analysis of multivariate curve resolution – alternating least square/discriminant analysis (MCR-ALS/DA) resulted in accurate and reliable results. Better results for print identifications were presented in UV range (83%) compared to ranges of Vis (60%) and UV-Vis (71%).	[66]

Lower standard deviation values were obtained using datasets from μ XRF analysis (approximately 20%) in comparison to LIBS (approximately 28%) when taking into consideration the quantitative repeatability of these methods [56]. Therefore, the repeatability of LIBS is slightly lower than μ XRF. Principal Component Analysis (PCA) was applied for toner and inkjet ink samples. It was found that PCA offered good separation and variation (beyond 95%) between brands and models of these samples [1,2]. PCA required an extensive sampling collection, as highlighted by Trejos *et al.* [7]. On top of the sampling requirement, a systematic grouping of printing ink based on their source and brand requires the necessary information from the manufacturers of the ink.

Haase *et al.* [62] implemented PCA prior to classifications using LDA and KNN. PCA reduced the variables to be used as data input for these supervised machine-learning models. The classifications with KNN and LDA were realised by combining the data from ATR-FTIR and LIBS. The hybrid information produced robust classification models with low error and confusion rates. These models had high accuracy, as observed in the classification rate for both LDA (85% - 99%) and KNN (90% - 100%). In contrast to models from separate datasets, lower percentages were observed in LIBS datasets, with 69% to 94% for LDA and 74% to 97% for KNN. ATR-FTIR datasets gave outputs of 81% to 95% and 88% to 99% for KNN and LDA, respectively. Excellent discrimination was achieved for toner (97% with five indistinguishable pairs) and inkjet (93% with one indistinguishable pair) samples on paper when the 3-sigma criterion was employed [1].

Applying a pairwise comparison analysis namely ANOVA with Tukey's post hoc test efficiently discriminated toner and inkjet ink samples with 98% and 100% discrimination, respectively. The three indistinguishable pairs of toner ink samples were discriminated using student t-test at a 95% confidence limit. The technique of MCR-ALS/DA acquired reliable and accurate results, as conducted by Metzinger *et al.* [66]. Better results were observed in the UV range for print identifications compared to ranges of Vis and UV-Vis. This is proved by a higher percentage of accuracy in the UV range at 83% with 40 out of 48 correct identifications. As for the rest of the spectral ranges, the Vis range had an accuracy of 60%, while the UV-Vis range was accurate by 71% for print identifications. UV range showed low spectral background and number of peaks resulting in highly distinctive spectral features. Therefore, better results were observed in MCR-ALS/DA approach by using the UV range.

7. Conclusions

Analysing printing ink may be more complex than pen ink because of the broader variability of compounds that are present in the printing ink. The implementation of spectroscopy analysis, including LIBS, has garnered attention among forensic scientists. This review article discusses the strengths and limitations of this technique in the analysis of printing ink. Although LIBS' applications are still in their infancy, the reliability of LIBS to distinguish printing ink samples qualitatively and quantitatively has been peer reviewed in a few studies. LIBS provides

essential information on printing ink to support the results of organic constituents from ATR-FTIR spectroscopy while overcoming the analytical limitations of other techniques (LA-ICP-MS). Furthermore, LIBS showed promising results when coupled with statistical analysis that yields higher accuracy for ink discrimination. Future research can be narrowed to the application of LIBS in conjunction with spectroscopic and statistical techniques that have not been previously explored. The application of LIBS to analyse other questionable printed documents, such as travel and identity documents, can be explored for good measure.

Acknowledgement

The authors would like to acknowledge Universiti Teknologi MARA (UiTM), Malaysia for funding under the Geran Penyelidikan Khas (600-RMC/GPK 5/3 (235/2020)) as well as individuals who contributed to the content of this review article.

References

- [1] Lennard, Chris, Moteaa M. El-Deftar, and James Robertson. "Forensic Application of Laser-Induced Breakdown Spectroscopy for the Discrimination of Questioned Documents." *Forensic Science International* 254, (2015): 68–79.
<https://doi.org/10.1016/j.forsciint.2015.07.003>.
- [2] Hui, Yew Wan, Naji Arafat Mahat, Dzulkiflee Ismail, and Raja Kamarulzaman Raja Ibrahim. "Laser-Induced Breakdown Spectroscopy (LIBS) for Printing Ink Analysis Coupled with Principle Component Analysis (PCA)." *AIP Conference Proceedings* 2155, no. September (2019): 020010-1-020010–19.
<https://doi.org/10.1063/1.5125514>.
- [3] Kula, Agnieszka, Renata Wietecha-Posłuszny, Katarzyna Pasioneck, Małgorzata Król, Michał Woźniakiewicz, and Paweł Kościelniak. "Application of Laser Induced Breakdown Spectroscopy to Examination of Writing Inks for Forensic Purposes." *Science and Justice* 54, no. 2 (2014): 118–25.
<https://doi.org/10.1016/j.scijus.2013.09.008>.
- [4] Materazzi, Stefano, Roberta Risoluti, Sara Pinci, and Francesco Saverio Romolo. "New Insights in Forensic Chemistry: NIR/Chemometrics Analysis of Toners for Questioned Documents Examination." *Talanta* 174, (2017): 673–78.
<https://doi.org/10.1016/j.talanta.2017.06.044>.
- [5] Sharaa, S. I., A. A. Aboul El-Magd, A. A. Bakr, Y. M. Moustafa, A. A. Shabana, and I. M. Abd El-Aziz. "Some of the Physical and Chemical Characterizations Applied for the Laser Printers Toner and Ballpoint Pen Inks to Determine the Sequence of Their Intersections." *Egyptian Journal of Chemistry* 62, no. 11 (2019a): 2047–2060.
<https://doi.org/10.21608/EJCHEM.2019.10960.1702>.
- [6] Trejos, Tatiana, Alejandra Flores, and José R. Almirall. "Micro-Spectrochemical Analysis of Document Paper and Gel Inks by Laser Ablation Inductively Coupled Plasma Mass Spectrometry and Laser Induced Breakdown Spectroscopy." *Spectrochimica Acta - Part B Atomic Spectroscopy* 65, no. 11 (2010): 884–895.
<https://doi.org/10.1016/j.sab.2010.08.004>.
- [7] Trejos, Tatiana, Ruthmara Corzo, Kiran Subedi, and José Almirall. "Characterization of Toners and Inkjets by Laser Ablation Spectrochemical Methods and Scanning Electron Microscopy-Energy Dispersive X-Ray Spectroscopy." *Spectrochimica Acta - Part B Atomic Spectroscopy* 92, (2014): 9–22.
<https://doi.org/10.1016/j.sab.2013.11.004>.
- [8] Hark, Richard R., and Lucille J. East. (2014). *Forensic Applications of LIBS*. Springer

- Series in Optical Sciences*. Vol. 182.
https://doi.org/10.1007/978-3-642-45085-3_14.
- [9] Al Balah, Osama, and Olodia Aied Taha Nassef. "A Further Analysis of Laser Induced Breakdown Spectroscopy Ink Pens' Spectra Using Principal Component Analysis (PCA) for Forensic Characterization." *Arab Journal of Nuclear Sciences and Applications* 52, no. 2 (2019): 72–78.
<https://doi.org/10.21608/ajnsa.2019.4231.1097>.
- [10] "Analysis of Printing Inks - John B. Tuttle, W. Harold Smith - Google Books." (n.d.) Accessed May 28, 2021.
<https://books.google.com.my/books?id=l62D49iJk9sC&pg=PA4&dq=first+printing+ink+analysis&hl=en&sa=X&ved=2ahUKEwiWwPfOiOnwAhURPXAKHZATCVIQ6AEwAHoECACQAg#v=onepage&q=first printing ink analysis&f=false>.
- [11] Bellis, Mary. (2019). "History of Computer Printers."
<https://www.thoughtco.com/history-of-computer-printers-4071175>.
- [12] Buczynski, Ludwik. "Special Print Quality Problems of Ink Jet Printers." (1999).
- [13] Leon, D. A., P. Thomas, and S. Hutchings. "Lung Cancer among Newspaper Printers Exposed to Ink Mist: A Study of Trade Union Members in Manchester, England." *Occupational and Environmental Medicine* 51, no. 2 (1994): 87–94.
<https://doi.org/10.1136/oem.51.2.87>.
- [14] Tang, H. K., W. H. Kuek, Rita Gozali, and Rolan Liew. "Ink Jet Printer Pioneers and Competitors." *Proceedings of the 2000 IEEE International Conference on Management of Innovation and Technology* 1, (2000): 94–98.
<https://doi.org/10.1109/ICMIT.2000.917286>.
- [15] Han, Kyuseo, Hongki Kim, Byungho Kang, Changrak Yoon, Jinseo Kim, and Maengsub Cho. "Characterizing the Desktop Color Printer with Polynomial Regression." *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics* 5, (1998): 4369–4372. IEEE.
<https://doi.org/10.1109/icsmc.1998.727535>.
- [16] Pollock, Doug, and Remi Coulon. "Life Cycle Assessment of an Inkjet Print Cartridge." *IEEE International Symposium on Electronics & the Environment*, (1996): 154–160.
<https://doi.org/10.1109/isee.1996.501870>.
- [17] Hoehse, Marek, Andrea Paul, Igor Gornushkin, and Ulrich Panne. "Multivariate Classification of Pigments and Inks Using Combined Raman Spectroscopy and LIBS." *Analytical and Bioanalytical Chemistry* 402, no. 4 (2012): 1443–1450.
<https://doi.org/10.1007/s00216-011-5287-6>.
- [18] Menzyk, Alicja, Grzegorz Zadora, and Mieczysław Sajewicz. "Physicochemical Analysis of Ink - Dating and Establishing the Sequence of Intersecting Lines of Ink Entries." *Problems of Forensic Sciences* 104, (2015): 279–302.
- [19] Shaffer, Douglas K. "Forensic Document Analysis Using Scanning Microscopy." *Scanning Microscopy* 7378, (2009): 73782E.
<https://doi.org/10.1117/12.825186>.
- [20] Calcerrada, Matías, and Carmen García-Ruiz. "Analysis of Questioned Documents: A Review." *Analytica Chimica Acta* 853, no. 1 (2015): 143–166.
<https://doi.org/10.1016/j.aca.2014.10.057>.
- [21] Poon, N. L., S. S.H. Ho, and C. K. Li. "Differentiation of Coloured Inks of Inkjet Printer Cartridges by Thin Layer Chromatography and High Performance Liquid Chromatography." *Science and Justice - Journal of the Forensic Science Society* 45, no. 4 (2005): 187–194.

- [https://doi.org/10.1016/S1355-0306\(05\)71665-8](https://doi.org/10.1016/S1355-0306(05)71665-8).
- [22] Szafarska, Małgorzata, Renata Wietecha-Posłuszny, Michał Woźniakiewicz, and Paweł Kościelniak. "Examination of Colour Inkjet Printing Inks by Capillary Electrophoresis." *Talanta* 84, no. 5 (2011): 1234–1243.
<https://doi.org/10.1016/j.talanta.2010.12.024>.
- [23] Król, Małgorzata, Agnieszka Kula, Renata Wietecha-Posłuszny, Michał Woźniakiewicz, and Paweł Kościelniak. "Examination of Black Inkjet Printing Inks by Capillary Electrophoresis." *Talanta* 96, (2012): 236–242.
<https://doi.org/10.1016/j.talanta.2011.12.025>.
- [24] Król, Małgorzata, Agnes Karoly, and Paweł Kościelniak. "Raman Spectroscopy and Capillary Electrophoresis Applied to Forensic Colour Inkjet Printer Inks Analysis." *Forensic Science International* 242, (2014): 142–149.
<https://doi.org/10.1016/j.forsciint.2014.06.031>.
- [25] Trejos, Tatiana, Peter Torriano, Ruthmara Corzo, Ana Raeva, Kiran Subedi, Rhett Williamson, Jong Yoo, and Jose Almirall. "A Novel Forensic Tool for the Characterization and Comparison of Printing Ink Evidence: Development and Evaluation of a Searchable Database Using Data Fusion of Spectrochemical Methods." *Journal of Forensic Sciences* 61, no. 3 (2016): 715–724.
<https://doi.org/10.1111/1556-4029.13109>.
- [26] Szykowska, M. I., Czerski, K., Paryczak, T. and Parczewski, A. "Ablative Analysis of Black and Colored Toners Using LA-ICP-TOF-MS for the Forensic Discrimination of Photocopy and Printer Toners." *Surface and Interface Analysis* 42, no. 5 (2010): 429–437.
<https://doi.org/10.1002/sia.3194>.
- [27] Lee, Jihye, Yun Sik Nam, Jisook Min, Kang Bong Lee, and Yeonhee Lee. "TOF-SIMS Analysis of Red Color Inks of Writing and Printing Tools on Questioned Documents." *Journal of Forensic Sciences* 61, no. 3 (2016): 815–822.
<https://doi.org/10.1111/1556-4029.13047>.
- [28] Egan, William J., Randolph C. Galipo, Brian K. Kochanowski, Stephen L. Morgan, Edward G. Bartick, Mark L. Miller, Dennis C. Ward, and Robert F. Mothershead. "Forensic Discrimination of Photocopy and Printer Toners. III. Multivariate Statistics Applied to Scanning Electron Microscopy and Pyrolysis Gas Chromatography/Mass Spectrometry." *Analytical and Bioanalytical Chemistry* 376, no. 8 (2003): 1286–1297.
<https://doi.org/10.1007/s00216-003-2099-3>.
- [29] Melendez-Perez, Jose Javier, Deleon Nascimento Correa, Vinicius Veri Hernandez, Damila Rodrigues De Moraes, Rodrigo Borges De Oliveira, Wanderley De Souza, Jandyson Machado Santos, and Marcos Nogueira Eberlin. "Forensic Application of X-Ray Fluorescence Spectroscopy for the Discrimination of Authentic and Counterfeit Revenue Stamps." *Applied Spectroscopy* 70, no. 11 (2016): 1910–1915.
<https://doi.org/10.1177/0003702816645352>.
- [30] Sharaa, S. I., A. A. Aboul El-Magd, A. A. Bakr, Y. M. Moustafa, A. A. Shabana, and I. M. Abd El-Aziz. "The Physical Application of Non-Destructive Techniques in Detection the Sequence of Intersecting Gel Ink and Printed Laser Toner Strokes." *Egyptian Journal of Chemistry* 62, no. 6 (2019b): 1069–1091.
<https://doi.org/10.21608/EJCHEM.2019.6532.1549>.
- [31] Shara, S.I., Y.M. Moustafa, A.A. Bakr, A.A. Aboul El-Magd, and I.M. Abd El-Aziz. "Application of Some Physical Techniques for Forensic Discrimination of Printer Toner." *Egyptian Journal of Chemistry* 61, no. 1 (2018): 131–142.

- <https://doi.org/https://doi.org/10.21608/ejchem.2017.1545.1123>.
- [32] Verma, Neha, Raj Kumar, and Vishal Sharma. "Analysis of Laser Printer and Photocopier Toners by Spectral Properties and Chemometrics." *Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy* 196, no. February (2018): 40–48. <https://doi.org/10.1016/j.saa.2018.02.001>.
- [33] Corzo, Ruthmara, Kiran Subedi, Tatiana Trejos, and José R. Almirall. "Evaluation of the Forensic Utility of Scanning Electron Microscopy-Energy Dispersive Spectroscopy and Laser Ablation-Inductively Coupled Plasma-Mass Spectrometry for Printing Ink Examinations." *Journal of Forensic Sciences* 61, no. 3 (2016): 725–734. <https://doi.org/10.1111/1556-4029.13110>.
- [34] Kumar, Kapil, and Shobhana Menon. "Determination of Relative Age of Inkjet and Laser Printouts by CIE L * a * b * Color Space." *International Journal of Sciences: Basic and Applied Research (IJSBAR)* 31, no. 2 (2017): 226–236.
- [35] Williamson, Rhett, Anna Raeva, and Jose R. Almirall. "Characterization of Printing Inks Using DART-Q-TOF-MS and Attenuated Total Reflectance (ATR) FTIR." *Journal of Forensic Sciences* 61, no. 3 (2016): 706–714. <https://doi.org/10.1111/1556-4029.13107>.
- [36] Verma, Neha, Vishal Sharma, Raj Kumar, R. Sharma, M. C. Joshi, G. R. Umapathy, Sunil Ohja, and Sundeep Chopra. "On the Spectroscopic Examination of Printed Documents by Using a Field Emission Scanning Electron Microscope with Energy-Dispersive X-Ray Spectroscopy (FE-SEM-EDS) and Chemometric Methods: Application in Forensic Science." *Analytical and Bioanalytical Chemistry* 411, no. 16 (2019): 3477–3495. <https://doi.org/10.1007/s00216-019-01824-z>.
- [37] Moore, Katherine Louise, Marko Barac, Marko Brajković, Melanie Jane Bailey, Zdravko Siketić, and Iva Bogdanović Radović. "Determination of Deposition Order of Toners, Inkjet Inks and Blue Ballpoint Pen Combining MeV Secondary Ion Mass Spectrometry and Particle Induced X-Ray Emission." *Analytical Chemistry* 91, no. 20 (2019): 12997–13005. <https://doi.org/10.1021/acs.analchem.9b03058>.
- [38] Brech F, Cross L. "Optical microemission stimulated by a ruby maser." *Applied Spectroscopy* 16, (1962): 59.
- [39] Kaiser, Jozef, Novotný, Karel, Martin, Madhavi Z, Hrdlička, Aleš, Malina, Radomír, Hartl, Martin, Adam, Vojtěch and Kizek René. "Trace elemental analysis by laser-induced breakdown spectroscopy—biological applications." *Surface Science Reports* 67, (2012): 233-243. <https://doi.org/10.1016/j.surfrep.2012.09.001>.
- [40] Rehse, Steven J, Salimnia H, Miziolek AW. "Laser-induced breakdown spectroscopy (LIBS): an overview of recent progress and future potential for biomedical applications." *Journal of Medical Engineering Technology* 36, (2012): 77-89. <https://doi.org/10.3109/03091902.2011.645946>
- [41] Santos Jr., Dário, Nunes, Lidiane Cristina, de Carvalho, Gabriel Gustinelli Arantes, Gomes, Marcos da Silva, de Souza, Paulino Florêncio, Leme, Flavio de Oliveira, dos Santos, Luis Gustavo Cofani, and Krug, Francisco José. "Laser-induced breakdown spectroscopy for analysis of plant materials: a review." *Spectrochimica Acta B Atomic Spectroscopy* 71-72, (2012): 3-13. <https://doi.org/10.1016/j.sab.2012.05.005>.
- [42] Pořízka, Pavel, Jakub Klus, Erik Képeš, David Prochazka, David W Hahn, and Jozef

- Kaiser. "On the Utilization of Principal Component Analysis in Laser-Induced Breakdown Spectroscopy Data Analysis , a Review." *Spectrochimica Acta Part B* 148 (2018): 65–82.
<https://doi.org/10.1016/j.sab.2018.05.030>.
- [43] Morel, Stéphane, Leone, Nicolas, Adam, Philippe, and Amouroux Jacques. "Detection of bacteria by time-resolved laser-induced breakdown spectroscopy." *Applied Optics* 42, no. 30 (2003): 6184-6191.
<https://doi.org/10.1364/AO.42.006184>.
- [44] Harmon, Russell S., Russo, Richard E., and Hark, Richard R. "Applications of laser-induced breakdown spectroscopy for geochemical and environmental analysis: a comprehensive review." *Spectrochimica Acta B Atomic Spectroscopy* 87, (2013): 11-26.
<https://doi.org/10.1016/j.sab.2013.05.017>.
- [45] Corsi, Michela, Cristoforetti, Gabriele, Hidalgo, Montserrat, Legnaioli, Stefano, Palleschi, Vincenzo, Salvetti, Azenio, Tognoni, Elisabetta, and Allebona Chiara. "Application of laser-induced breakdown spectroscopy technique to hair tissue mineral analysis." *Applied Optics* 42, no. 30 (2003): 6133-6137.
<https://doi.org/10.1364/AO.42.006133>.
- [46] Dolgin, Bella, Chen, Yuheng, Bulatov, Valery, and Schechter, Israel. "Use of LIBS for rapid characterization of parchment." *Analytical and Bioanalytical Chemistry* 386, no. 5 (2006): 1535-1541.
<https://doi.org/10.1007/s00216-006-0676-y>.
- [47] Melessanaki, K., Papadakis, V., Balas, C., and Anglos, D.. "Laser Induced Breakdown Spectroscopy and Hyper-Spectral Imaging Analysis of Pigments on an Illuminated Manuscript." *Spectrochimica Acta Part B* 56, no. 12 (2001): 2337–2346.
- [48] Ahmed, Harby E. and Nassef, Olodia Aied. "From Ptolemaic to modern inked linen via laser induced breakdown spectroscopy (LIBS)." *Analytical Methods* 5, no. 12 (2013): 3114-3121.
<https://doi.org/10.1039/C3AY26358D>.
- [49] Ponterio, R., Trusso, S., Vasi, C., La Torre, G. F., and Toscano Raffa, A. "Laser induced breakdown spectroscopy for the analysis of archaeological dyes from Licata (Sicily)." *Radiation Effects and Defects in Solids* 163, (2008): 535-543.
<https://doi.org/10.1080/10420150701780755>.
- [50] Noll R. (2012). *Laser-induced Breakdown Spectroscopy Fundamentals and Applications*. 1st edn. Heidelberg: Springer Berlin.
- [51] Yun, Jong-Il, Klenze, Reinhardt, and Kim, Jae-Il. "Laser-Induced Breakdown Spectroscopy for the On-Line Multielement Analysis of Highly Radioactive Glass Melt. Part I: Characterization and Evaluation of the Method." *Applied Spectroscopy* 56, no. 4 (2002): 437-448.
<https://doi.org/10.1366/0003702021955097>.
- [52] Oujja, M., Vila, A., Rebollar, E., García, J. F., and Castillejo, M. "Identification of Inks and Structural Characterization of Contemporary Artistic Prints by Laser-Induced Breakdown Spectroscopy." *Spectrochimica Acta Part B: Atomic Spectroscopy* 60, no. 7–8 (2005): 1140–1148.
<https://doi.org/10.1016/j.sab.2005.05.021>.
- [53] Melessanaki, Kristalia, Maripaz Mateo, Susan C. Ferrence, Philip P. Betancourt, and Demetrios Anglos. "The Application of LIBS for the Analysis of Archaeological Ceramic and Metal Artifacts." In *Applied Surface Science*, 197–198. p. 156–63. 2002. Elsevier.

- [https://doi.org/10.1016/S0169-4332\(02\)00459-2](https://doi.org/10.1016/S0169-4332(02)00459-2).
- [54] Mateo, Mari Paz, Ctvrtnickova, Tereza, and Nicolas Gines. "Characterization of pigments used in painting by means of laser-induced plasma and attenuated total reflectance FTIR spectroscopy." *Applied Surface Science* 255, (2009): 5172–5176.
 - [55] Pořizka P, Prochazková P, Prochazka D, Sládková L, Novotný J, Petrilak M, Kaiser J. "Algal biomass analysis by laser-based analytical techniques—a review." *Sensors* 14, no. 12 (2014): 17725–17752.
 - [56] Król, Małgorzata, Klaudia Gondko, Agnieszka Kula, Patryk Własiuk, Julio M. del Hoyo-Meléndez, and Paweł Kościelniak. "Characterization of the Elemental Composition of Polish Banknotes by X-Ray Fluorescence and Laser-Induced Breakdown Spectroscopy." *Spectrochimica Acta - Part B Atomic Spectroscopy* 169, (2020): 105898.
<https://doi.org/10.1016/j.sab.2020.105898>.
 - [57] Almirall, Jose. "Laser Ablation Inductively Coupled Plasma Mass Spectrometry (LA-ICP-MS) and Laser Induced Breakdown Spectroscopy (LIBS) Analyses of Paper, Inks, and Soils," (2017):1–14.
 - [58] Zhang, Tianlong, Hongsheng Tang, and Hua Li. "Chemometrics in Laser-Induced Breakdown Spectroscopy." *Journal of Chemometrics* 32, no. 11 (2018): 1–18.
<https://doi.org/10.1002/cem.2983>.
 - [59] Spizzichino, Valeria, and Roberta Fantoni. "Laser Induced Breakdown Spectroscopy in Archeometry : A Review of Its Application and Future Perspectives ☆." *Spectrochimica Acta Part B: Atomic Spectroscopy* 99, (2014): 201–209.
<https://doi.org/10.1016/j.sab.2014.07.003>.
 - [60] Ikezawa, Satoshi, Muneaki Wakamatsu, Joanna Pawlat, and Toshitsugu Ueda. 2011. *Multi-Spectral Analytical Systems Using LIBS and LII Techniques*. Edited by Subhas Chandra Mukhopadhyay, Aimé Lay-Ekuakille, and Anton Fuchs. *New Developments and Applications in Sensing Technology*. 2011th ed. Vol. 83. Springer.
 - [61] Anabitarte, F., A. Cobo, and J. M. Lopez-Higuera. 2012. "Laser-Induced Breakdown Spectroscopy: Fundamentals, Applications, and Challenges." *International Scholarly Research Notices* 2012, (2012): 1–12.
<https://doi.org/10.5402/2012/285240>.
 - [62] Haase, Emily, Luis Arroyo, and Tatiana, Trejos. "Classification of Printing Inks in Pharmaceutical Packages by Laser-Induced Breakdown Spectroscopy and Attenuated Total Reflectance-Fourier Transform Infrared Spectroscopy." *Spectrochimica Acta - Part B Atomic Spectroscopy* 172, (2020): 105963.
<https://doi.org/10.1016/j.sab.2020.105963>.
 - [63] Elsherbiny, Nany and O Aied Nassef. "Wavelength Dependence of Laser Induced Breakdown Spectroscopy (LIBS) on Questioned Document Investigation." *Science & Justice* 55, no. 4 (2015): 254–263.
<https://doi.org/10.1016/j.scijus.2015.02.002>.
 - [64] Jiao, Wei, Zhang Tianlong, Dong Juan, Sheng Liwen, and Tang Hongsheng. "Quantitative Determination of Cr in Ink by Laser-Induced Breakdown Spectroscopy (LIBS) Using ZnO as Adsorbent." *Chemical Research in Chinese Universities* 31, (2015): 909–913.
<https://doi.org/10.1007/s40242-015-5210-3>.
 - [65] National Institute of Standards and Technology [Internet], [28 May 2021] Available from: <https://www.nist.gov/>.
 - [66] Metzinger, Anikó, Róbert Rajkó, and Gábor Galbács. "Discrimination of Paper and Print Types Based on Their Laser Induced Breakdown Spectra." *Spectrochimica Acta - Part*

- B Atomic Spectroscopy* 94–95, (2014): 48–57.
<https://doi.org/10.1016/j.sab.2014.03.006>.
- [67] Subedi, Kiran, Tatiana Trejos, and José Almirall. “Forensic Analysis of Printing Inks Using Tandem Laser Induced Breakdown Spectroscopy and Laser Ablation Inductively Coupled Plasma Mass Spectrometry.” *Spectrochimica Acta - Part B Atomic Spectroscopy* 103–104, (2015): 76–83.
<https://doi.org/10.1016/j.sab.2014.11.011>.
- [68] Martinez-Lopez, Claudia, Masataka Sakayanagi, and Jose R. Almirall. “Elemental Analysis of Packaging Tapes by LA-ICP-MS and LIBS.” *Forensic Chemistry* 8, (2018): 40–48.
<https://doi.org/10.1016/j.forc.2018.01.004>.
- [69] Tian-long, Zhang, Shan, Wu, Hong-sheng, Tang, Wang, Kang, Yi-xiang, Duan and Hua, Li. “Progress of Chemometrics in Laser-Induced Breakdown Spectroscopy Analysis.” *Chinese Journal of Analytical Chemistry* 43, no. 6 (2015): 939–948.
[https://doi.org/10.1016/S1872-2040\(15\)60832-5](https://doi.org/10.1016/S1872-2040(15)60832-5).
- [70] U.S. Geological Survey, Minerals Yearbook, 2008, V. 1, Metals and Minerals, U.S. Government Printing Office, 2011.
- [71] Iftime, Gabriel, Vanbesien, Daryl W., Birau, Maria, Wosnick, Jordan, Kazmaier, Peter M, Toner Containing Fluorescent Nanoparticles, Google Patents, 2013.
- [72] Rzecki, Krzysztof, Tomasz Sośnicki, Mateusz Baran, Michał Niedźwiecki, Małgorzata Król, Tomasz Łojewski, U. Rajendra Acharya, Özal Yildirim, and Paweł Pławiak. “Application of Computational Intelligence Methods for the Automated Identification of Paper-Ink Samples Based on LIBS.” *Sensors* 18, no. 11 (2018): 3670.
<https://doi.org/10.3390/S18113670>.

e ISBN 978-967-2426-49-3



Published By:



GLOBAL ACADEMIC EXCELLENCE (M) SDN BHD -1257579-U

Lot 2 -11 Arked MARA Kota Bharu, Jalan Dato Pati, 15000 Kota Bharu ,Kelantan, MALAYSIA

Website: www.egax.org

Email: admin@egax.org