

ISSN 3031-5514



PROCEEDING

THE 6TH INTERNATIONAL CONFERENCE ON SCIENCE AND TECHNOLOGY (ICST)

SEMIRATA BKS PTS WILAYAH
BARAT BIDANG FMIPA 2023

Banjarbaru,
South Kalimantan, Indonesia
on July 7-8th, 2023



*“Harmonizing People, Nature, and Culture Through
Scientific Collaboration”*

FACULTY OF MATHEMATICS AND NATURAL SCIENCES
LAMBUNG MANGKURAT UNIVERSITY

Organized by



PROCEEDINGS

The 6th International Conference of Sciences and Technology (ICST)

*“Harmonizing People, Nature, and Culture Through
Scientific Collaboration”*

Volume 1, 2023

Banjarbaru, 07 – 08 July 2023

Publisher



Lambung Mangkurat University Press

ORGANIZING COMMITTEES

Advisory Board

1. Prof. Dr. Ahmad, SE., M.Si. (Rektor)
2. Prof. Drs. Abdul Gafur, M.Si, M.Sc, Ph.D. (Dekan FMIPA)
3. Dr. Gunawan, S.Si., M.Si.
4. Dr. Totok Wianto, S.Si., M.Si.
5. Mohammad Reza Faisal, ST., MT., Ph.D.

Scientific Committee

Chief of Executive	: Prof. Dewi Anggraini, S.Si., M.App.Sci., Ph.D* Rani Sasmita, S.Si., M.P., M.Sc.
General Programme	: Dewi Sri Susanti, S.Si., M.Si.* Dr. Dra. Ninis Hadi Haryanti, M.Si Dr. Drs. Heri Budi Santoso, M.Si Dr. Uripito Trisno Santos, S.Si., M.Si
Secretary Section	: Utami Irawati, S.Si, M.ES., Ph.D* Dr. Na'imah Hijriati, S.Si., M.Si Sasi Gendro Sari, S.Si., M.Sc Gusti Atika Urfa, S.Si., M.Sc. Aprida Siska Lestia, S.Si., M.Si Rahmat Eko Sanjaya, M.Si. Zuhra Zuraida, S.H Razmeirahmini, S.E Indah Purwaningasti
Finance, Sponsorship, and Accommodation Section	: Sri Cahyo Wahyono, S.Si., M.Si.* Oni Susanto, S.Si., M.Si Maisarah, SE. Dhian Asmarini S.Si.
Programme Section	: Dr.Nurma Sari, S.SI., M.Si.* Dr. Nurlina, S.Si., M.Sc Saman Abdurrahman, S.Si., M.Sc. Yuana Sukmawaty, S.Si., M.Si apt. Fadlilaturahmah, S.Farm., M.Sc apt. Mia Firiana, S.Farm., M.Si Simon Sadok Siregar, S.Si., M.Si apt. Satrio Wibowo Rahmatullah, S.Farm., M.Sc apt. Deni Setiawan, S.Farm., Mclin.Pharm Rudy Herteno, S.Kom., M.Kom apt. Dita Ayulia Dwi Sandi, S.Farm., M.Sc

	Selvi Annisa, S.Si., M.Si Hermei Lissa, S.Pd., M.Si
Publication Section	: Dr. Muhammad Ahsar Karim, S.Si., M.Sc.* Yuni Yulida, S.Si., M.Sc. Dr. Moch Idris, S.Si., M.Si Prof. Rodiansono, S.Si., M.Si., Ph.D Maria Dewi Astuti, S.Si., M.Si Dewi Umaningrum, S.Si., M.Si Dr. Drs. Krisdianto, M.Sc Anang Kadarsah, S.Si., M.Si Dr. Tetti Novalina Manik, S.Si., M.T Ade Agung Harnawan, S.Si., M.Sc Dr. apt. Sutomo, S.Si., M.Si apt. Muhammad Ikhwan Rizki, S.Farm., M.Farm Muliadi, S.Kom., M.Cs Fuad Muhajirin Farid, S.Pd., M.Si.
General Treasure, IT, Recording, and Documentation Section	: Dodon Turianto Nugrahadi, S.Kom., M.Eng.* Muhammad Itqan Mazdadi, S.Kom., M.Kom. Triando Hamonangan Saragih S.Kom., M.Kom. Syaipullah, S.Mn Yodia Rossanto Tirto, S.Si Muhammad Sholih ‘Afif, S.Kom
Competition Section	: apt. Prima Happy Ratnapuri, S.Farm., M.Sc.* Noer Komari, S.Si, M.Kes Nur Faridah, S.E apt. Okta Muthia Sari, M.Farm apt. Hayatun Izma, M.Farm.Sci Yeni Rahkmawati, S.Mat., M.Si Rizki Fitria M.Si Aida Fitri, S.E Yuyun Magfirah, S.E
Field Trip Section	: apt. Nani Kartinah, S.Farm., M.Sc.* apt. Adtya Maulana Perdana Putra, S.Farm., M.Sc. Nie’matul Ridha S.E Fitri Dwiyani, S.Si.

EDITORIAL BOARD

Editor-in-Chief

Dr. Muhammad Ahsar Karim, M.Sc. (Universitas Lambung Mangkurat, Indonesia)

Deputy Editor

Yuni Yulida, M.Sc. (Universitas Lambung Mangkurat, Indonesia)

Associate Editors

Prof. Tim Roberts (The University of Newcastle, Australia)

Prof. Rodiansono, M.Si., Ph.D. (Universitas Lambung Mangkurat, Indonesia)

Prof. Setia Pramana, Ph.D. (Politeknik Statistika STIS, Indonesia)

Dr. Deden Rukmana (Alabama A & M University, United States)

Prof. Drs. Abdul Gafur, M.Si., M.Sc., Ph.D. (Universitas Lambung Mangkurat, Indonesia)

Prof. Dewi Anggraini, M.App.Sci., Ph.D. (Universitas Lambung Mangkurat, Indonesia)

Prof. Dr. apt. Sutomo, M.Si. (Universitas Lambung Mangkurat, Indonesia)

Bustang, S.Pd., M.Sc., Ph.D. (Makassar State University, Indonesia)

Dr. Tetti Novalina Manik, M.T. (Universitas Lambung Mangkurat, Indonesia)

Dr. Moch. Idris, M.Si. (Universitas Lambung Mangkurat, Indonesia)

Maria Dewi Astuti, M.Si. (Universitas Lambung Mangkurat, Indonesia)

Anang Kadarsah, M.Si. (Universitas Lambung Mangkurat, Indonesia)

Fuad Muhajirin Farid, M.Si. (Universitas Lambung Mangkurat, Indonesia)

apt. Muhammad Ikhwan Rizki, M.Farm. (Universitas Lambung Mangkurat, Indonesia)

Muliadi, M.Cs. (Universitas Lambung Mangkurat, Indonesia)

Reviewer:

Prof. Sunardi, M.Sc., Ph.D

Utami Irawati, Ph.D

Dr. Muhammad Ahsar K., M.Sc.

Dr. Na'imah Hijriati, M.Si.

Dr. Moch. Idris, M.Si

Dr. Samsul Hadi

Dr. Heri Budi Santoso, M.Si.

Amar Vijai Nasrulloh, Ph.D

Noer Komari, M. Kes

Rahmat Eko Sanjaya, M.Si

Kholifatu Rosyidah, M.Si

Sasi Gendro Sari, M.Sc.

Rizki Fitria, M.Si

Witiyasti Imaningsih, M.Si.

Simon Sadok Siregar, M.Si.

Selvi Annisa, M.Si.

Yeni Rahkmawati, M.Si.

Yuana Sukmawaty, M.Si.

Pardi Affandi, M.Sc.

Aprida Siska Lestia, M.Si.

Saman Abdurrahman, M.Sc

Muliadi, M.Cs.

Keynote Speaker:

Emeritus Professor Tim Roberts

School of Environmental and Life Sciences, The University of Newcastle – Australia

Dr. Deden Rukmana

Department of Community and Regional Planning, Alabama A&M University-USA

Prof. Rodiansono, S.Si., M.Si., Ph.D

Department of Chemistry, Faculty of Mathematics and Natural Sciences, Universitas Lambung Mangkurat – Indonesia

Invited Speakers:

Prof. Setia Pramana, S.Si., Ph.D.

Head of Data Processing and Analysis, Statistics Indonesia – Indonesia.

Computational Statistics Department, Politeknik Statistika STIS Jakarta, Indonesia.

Dr. Eng. Imam Tahyudin, S.Si., M.Kom., M.M.

Faculty of Computer Science, Amikom Purwokerto University - Indonesia

Dr. Ir. Badruzsaufari, M.Sc.

Department of Biology, Faculty of Mathematics and Natural Sciences, Universitas

Lambung Mangkurat - Indonesia

Dr. Sri Yani Purwaningsih, S.Si., M.Si.

Faculty of Science and Data Analytics, Department of Physics, Institut Teknologi

Sepuluh Nopember – Indonesia

Prof. Ejeatuluchukwu Obi, Ph.D.

Department of Pharmacology and Therapeutics, Faculty of Medicine, College of

Health Sciences, Nnamdi Azikiwe University - Nigeria

Bustang, S.Pd, M.Sc., Ph.D.

Department of Sport Science Faculty of Sport Science and Health Universitas Negeri

Makassar

Issue:

December 2023

ACKNOWLEDGMENTS

This proceeding is dedicated to keynote speakers, invited speakers and researchers who have been participated on The 6th International Conference of Sciences and Technology (ICST) with the theme of “Harmonizing People, Nature, and Culture Through Scientific Collaboration”, so that the conference can be carried out properly and this proceedings can be published.

The theme was chosen with the reason to pay attention to the academic world about the importance of harmonizing people, nature and culture through scientific collaboration. National academics have produced a lot of scientific research, both independently and collaboratively, on the harmonization of people, nature and culture. However, quite a few of these studies have not been widely disseminated or published, so that the research cannot be accessed by people who need it. On that basis, this international conference has become an event for national academics to present their research, as well as exchange information and deepen studies of their research problems, as well as develop sustainable collaboration.

There were 155 scientific articles submitted by the authors to the conference committees to be presented in parallel sessions. The authors come from various institutions, ranging from universities, research institutes, and practitioners. With the publication of this proceedings, it is hoped that the research results can be used as references in developing further related research. The editorial team would like to thank all parties involved in the completion of this proceedings.

Best Regard,

The 6th ICST Committees

TABLE OF CONTENTS

COVER	i
ORGANIZING COMMITTEES	ii
EDITORIAL BOARD	iv
ACKNOWLEDGMENTS	vi
TABLE OF CONTENTS	vii

Biology

Anang Kadarsah, Dini Sofarini, and Devy Azizah Rahman

Presence of Water Pollution Sources on Phytoplankton Abundance and Water Quality in Tamiyang Lake of Banjar Regency

.....1

Rina Falhiyah, Aulia Ajizah, and Amalia Rezeki

Variety of Ferns in Sultan Adam Grand Forest Park

.....7

Rochmah Agustina, Nismah Nukmal, C.N. Ekowati, and Gina Dania Pratami

Toxicity of Methanol Extract from Kitolod (*Hippobroma Longiflora* (L.) G. Don) Leaves and Stems to Mortality Rate of Cocoa Mealybug (*Planococcus Minor* Maskell., Hemiptera: Pseudococcidae)

.....18

Jani Master, Santoso, Ahmad Fanani, Sumianto, Ichan Prastika, and Muhammad Yunus

Analysis of Coverage and Distribution of *Melastoma malabathricum* Invasion in Way Kambas National Park

.....25

Rinta Dwi Takarini, Rani Sasmita, Badruzaufari

Morphological Variation of Climbing Perch (*Anabas testudineus* Bloch 1972) in South Kalimantan Based on Morphometric Characters

.....30

Yurnaliza, Nunuk Priyani, Irina Levinita

Screening of Phosphate-Solubilizing Fungi from Oil Palm Rhizosphere

.....37

Muhammad Syukri Fadil, Djong Hon Tjong, Zozy Aneloi Noli, Eti Yerizel

Effect of the Methanol Extract of Dragon Scales Leaves (*Drymoglossum piloselloides* [L.] Presl.) on Malondialdehyde and Catalase Level in Blood Serum of Lead-Exposed Albino Rats
42

Chemistry**Noer Komari, Eko Suhartono, Suhaili Asmawi**

Mercury Binding Antioxidant Enzymes of *Oreochromis Niloticus*: In Silico Evaluation
47

Musbahu Adam Ahmad, Sri Sumarsih, Siti Febtria Asrini Sugito, Aswandi Wibrianto, Mochamad Zakki Fahmi

Calotropis Gigantea Derived Carbon Dots For Breast Cancer Bioimaging
58

Refilda, Andriani Febriyanti, Yefrida

Nutrient Analysis of Compost from Vegetable and Fruit Waste and Its Application to Hydroponic *Ipomoea reptana* Plant
64

Nafisah, Hanna Habibah, Sarmila, Indah Saputri, Indah Setiawati and Noer Komari

Effects of Kelakai (*Stenochlaena Palutris*) Extract on Exposure to Organophosphate Pesticide: In Silico and In Ovo Cytotoxic Studies
73

Annisa Elcentia Fajarwati, Mutia Hardhiyuna, Febby Nurdiya Ningsih, Andi Setiawan, Ucu Yanu Arbi, Dwi Wahyu Indriani, and Peni Ahmadi

The Rapid Identification Metabolites of *Aaptos* sp. from Bunaken National Park, Manado
84

Azidi Irwan, Yuni Kurniawati, Kamilia Mustikasari

Analysis Of Limau Kuit Leaves Volatile Compounds Obtained By Soxhlet-N-Hexane Extraction With Gc-MS
92

Rahmat Gunawan, Fahmi Fadillah, Desmond Kristian, Chairul Saleh, and Erwin

Ruthenium Complex and Methyl Dendrimer Compounds as Artificial Photosynthesis.
First Principle Study for Renewable Energy

.....100

Physics

Minarni Shiddiq, Alvika Juliardi Pratama, Zikri Radikel, Nadia Zakyyah Yasmin

Laser Induced Fluorescence Imaging for Ripeness Detection of on-Tree-Oil Palm Fresh
Fruit Bunches

.....104

**Arif Ismul Hadi, M. Farid, Lindung Zalbuin Mase, Refrizon, Shela Basaria
Purba, and Darmawan Ikhlas Fadli**

Analysis of Soil Site Classes Based on Microtremor Measurements in the South
Bengkulu, Indonesia

.....113

Muhammad Farid, Arif Ismul Hadi, Fadillah Humairah

Investigation of Landslide in the Ulu Manna, South of Bengkulu District Based on
Horizontal to Vertical Spectral Ratio Inversion

.....121

Mathematics And Statistics

**Syaripuddin, Fidia Deny Tisna Amijaya, Wasono, Shanaz Tulzahrah, Rara
Suciati**

Application of Quadratic Programming Using Wolfe Method on Portfolio Optimization

.....129

Urai Fiqri Fahturrahman, Nilamsari Kusumastuti, and Bayu Prihandono

Homomorphic Ordering of the Max-plus Formal Polynomials and the Max-plus
Polynomial Functions

.....137

Widya Kusuma Arum, Arum Handini Primandari

Clustering Analysis of Lecturer Profiles Using Gower Distance and K-Medoids
Algorithm

.....147

Yulian Fauzi, Zulfia Memi Mayasari, and Agus Susanto

Numerical Simulation of the Tsunami Wave Propagation Model Using Lax-Friedrichs Method
 158

Mariatul Qibtiah and Pardi Affandi

Optimal Control of Dengue Fever with the Influence of Vaccination and Insecticide Administration
 163

Ramya Rachmawati

A Probabilistic Dynamic Programming Approach for Solving an Inventory Problem (Case study: Fertilizer Stock)
 174

Regina Dietya Erika, Nur Salam, Aprida Siska Lestia

Application of the Arima Method for Forecasting Coffee Production in South Kalimantan Province
 182

Rahma Dina Nur Azizah, Dewi Sri Susanti, Selvi Annisa

Predicting the Consumer Price Index in South Kalimantan Province Using Support Vector Regression Method
 190

Ratna Annisa Kumala Dewi, Nur Salam, Yeni Rahkmawati

Intervention Analysis in Forecasting the Number of Arrival Passangers at Syamsudin Noor Banjarbaru
 198

Joannes Fregis Philosovio Anugrahnu, Naomi Nesyana Debataraaja, Dadan Kusnandar

Synthetic Minority Oversampling Technique Implementation on Random Forest in Archipelagic Fishing Port of Pemangkat
 208

Sri Wahyuningsih, Meiliyani Siringoringo, Ika Purnamasari¹, Darnah A. Nohe, Ruli Yuniarto, Jumarni, and Siti Nur Rahmah

Forecasting Reference Coal Price Using Transfer Function Multiple Inputs
 213

Dyah Yunita Hartanti, Dian Agustina, Winalia Agwil, Sigit Nugroho

Spatial Structural Equation Modeling of Poverty Indicators in the South Sumatra
Region

.....222

Presence of Water Pollution Sources on Phytoplankton Abundance and Water Quality in Tamiyang Lake of Banjar Regency

Anang Kadarsah^{*(1)}, Dini Sofarini⁽²⁾, and Devy Azizah Rahman⁽³⁾

⁽¹⁾Biology Program Study, Universitas Lambung Mangkurat, Indonesia

⁽²⁾Water Resources Management, Universitas Lambung Mangkurat, Indonesia

⁽³⁾Biology Program Study, Universitas Lambung Mangkurat, Indonesia

*Corresponding Author: anangkadarsah@ulm.ac.id

Abstract. This study aims to describe the pollution sources from anthropogenic activity that affect the phytoplankton abundance and water quality in Tamiyang Lake, Banjar District, South Kalimantan. Collection of pollution sources using the transect method along the edge of the lake as far as 600 meters and calculated based on the area covered by anthropogenic activities. Water quality (pH, temperature, Dissolved Oxygen (DO), Nitrate, Phosphate and Biological Oxygen Demand (BOD)) is measured at three sources of water pollution (gold mining, floating net cages and tourism zones). The results showed there are eleven main sources of water pollution in Lake Tamiyang, i.e : floating net cages, rice fields, gold mining, tourism area, fish pond, crumb rubber, broiler chickens and floating latrines. The biggest contribution to changes in water quality comes from floating net cages (71.2%), then gold mining (6.7%) and tourist areas (5%). The highest abundance of phytoplankton came from sand mining (31,524 Cells/L), and the lowest was from floating net cages (24,939 Cells/L). The values of temperature, pH, nitrate and phosphate are still below the permissible water quality threshold. The value of pH about 6.1 – 6.21, temperature about 29.51 - 30.1 Celsius, DO about 2.21-3.34 mg/l, Nitrate levels about 0.01 mg/l, Phosphate about 0.13-0.18 mg/l and BOD about 14.21-66.27 mg/l. Based on Chi-Square Test, there is no difference between pH, temperature, DO, Nitrate, and Phosphate, except for BOD at a value of = 000. Contribution : there is no source of pollution dominates changes in water quality.

Keywords : anthropogenic, phytoplankton, pollution, source, water quality

INTRODUCTION

Water is an important component to support physiological and non-physiological processes for living things. Water functions to meet daily needs, be it recreation, agriculture, fisheries, mining, and so on. As the population increases, the demand for water will also continue to increase (Bahri *et al.*, 2019). Quality and quantity of water at all times is very important to note in the context of sustainable conservation of water services [2]. Availability of water also is very important for sustainable development, but the anthropogenic activities will trigger more and more sources of pollution in water bodies. The more densely the population uses the waters and the pollution makes it difficult to obtain clean water, as happened in the Tamiyang Lake Area (Bahri *et al.*, 2015). This lake is a source of raw water for residents of Banjar Regency and its surroundings by damming the Riam Kanan watershed that crosses the West Mandikapu area. This lake has an area of 5 hectares with a depth of ±12 m (Haryati *et al.*, 2019).

Common sources of water pollution in this lake are floating net cages, and sand mining, as well as other activities. Floating net cages have the potential to become a source of pollution through the practice of giving excessive amounts and doses of feed. Unconsumed feed sources and fish excretion results will be a source of organic pollutants in the waters [5]. This will lead to nutrient enrichment and accelerate eutrophication which is marked by the development of aquatic plants such as *Eichhornia crassipes*, *Azolla Sp.* and *Hydrilla verticillata* (Anas *et al.*, 2017).

It was noted that the condition of the waters of Tami Lake was classified as lightly polluted with fertility status ranging from oligotrophic to eutrophic. Total Nitrogen has a range of 0.006-0.570 and total Phosphate shows a range of 0.007-0.230 (Haryati *et al.*, 2019). In this case, the presence of phytoplankton is influenced by the physico-chemical quality of the waters, so that it can be used as a bioindicator of water quality (Sudinno *et al.*, 2015). Phytoplankton is able to interact directly with changes in physico-chemical and

nutrient elements in a waters (Khairunnisa *et al.*, 2017). Phytoplankton are tolerant and have different responses to changes in water quality. The abundance of phytoplankton is a biomonitoring for water quality which is closely related to the measurement of environmental factors (Awaludin *et al.*, 2015).

So far there has been no publication regarding the source of water pollution and its impact on the phytoplankton community and water quality of Lake Tamiyang. Its Important to study the sources of pollution from anthropogenic activity in order to mitigate its effect on phytoplankton abundance and water quality. The occurrence of pollution in the waters, in addition to harming health through decreasing the quality of drinking water, can also cause loss of aquatic ecosystem services through eutrophication because it causes fish death and the resulting negative effect on a decrease in nutrition [2]. This study aims to describe the pollution sources from anthropogenic activity, the phytoplankton abundance and water quality in Tamiyang Lake, Banjar District of South Kalimantan.

RESEARCH METHODS

This research was conducted for 4 months starting from October 2022 to January 2023. The research location was in the Riam Kanan watershed, especially in the Tamiyang Lake area, which is in Mandikapau Barat Village, Karang Intan District, Banjar Regency, South Kalimantan (Figure 1). This lake is located about 15 kilometers east of Banjarbaru City, as the capital of South Kalimantan Province. This lake is of significant value to the residents of Banjar Regency, Banjarmasin City and Banjarbaru City through the provision of raw water by the Regional Drinking Water Company.



Figure 1. Map of research locations in Tamiyang Lake, Banjar Regency of South Kalimantan

Land use patterns and pollutant sources around the lake are identified using survey methods and field observations [9]. Collection of pollution sources along the edge of the lake using the transect method as far as 600 meters. The sources of pollution is calculated based on the percentage of covered area by anthropogenic activities.

Phytoplankton samples were taken by direct sampling method by filtering 30 L of water taken from a depth of 0-50 cm, then filtered using a plankton net. The sample collected then transferred to a 100 ml sample bottle. All phytoplankton samples were preserved by adding 2-6 drops of 5% Lugol's acid solution for every 100 ml of sample [10]. Plankton samples were obtained from 100 liters of filtered water and concentrated to 30 ml. The water sample was placed in a white or dark film bottle, and 4% formalin added. Plankton identified by using Optica B310 binoculars microscope with 100 times magnification. 1 ml of the homogenized samples, dripped on Sedgewick Rafter and covered with cover glass, then identified and calculated the abundance. Identification of plankton refers to plankton identification book. The abundance of plankton calculated referring to method by [11], by formula :

$$N = \frac{1}{V_d} \times \frac{V_t}{V_s} \times n$$

Description: N = Plankton abundance (cell/m³ or ind/m³); V_d = volume of filtered water sample (m³); V_t = volume of filtered sample (ml); V_s = concentrate volume of Sedgewick Rafter Counting Cell (ml); n = number of observed plankton.

Measurements of water quality parameters from the site were carried out by handheld Multi-Parameter Water Quality Checker (WQC)-24. Water quality (pH, temperature are measured at three sources of water pollution (gold mining, floating net cages and tourism zones). Behind it, Dissolved Oxygen (DO), Nitrate, Phosphate and Biological Oxygen Demand (BOD) and its referring to SNI 6989.57-2008 [12].

The results of the observations were then analyzed descriptively. Data analysis was carried out qualitatively and quantitatively. Descriptive qualitative analysis uses pictures and tables to analyze the relationship between pollutant sources and the abundance of phytoplankton and water quality. Then analytical quantitative analysis using ANOVA to see differences in physical and chemical characteristics [9]

RESULTS AND DISCUSSION

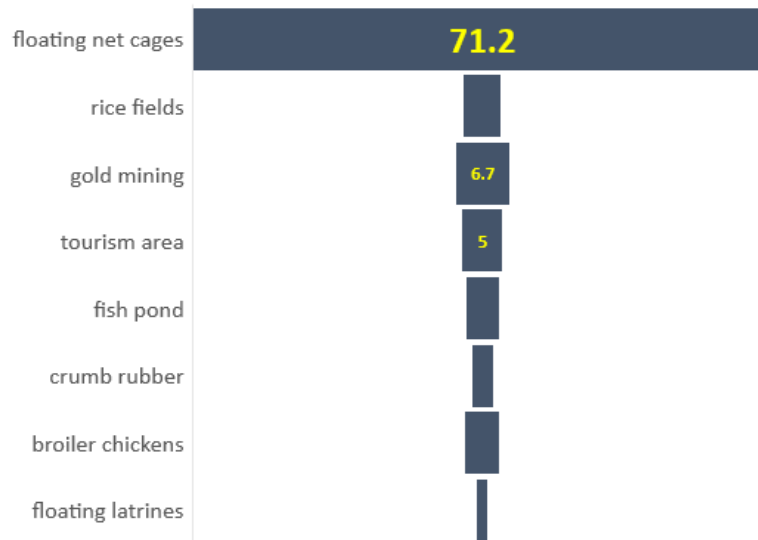
Main sources of water pollution

Lakes face various sources of pollution, impacting their water quality and ecological balance. Environmental and ecological degradation is more prominent in the lake especially in the highlands than in other ecosystems [13]. The results showed there are eleven main sources of water pollution in Lake Tamiyang, i.e : floating net cages, rice fields, gold mining, tourism area, fish pond, crumb rubber, broiler chickens and floating latrines. Figure 2 below shows four examples of sources of water pollution found in Lake Tamiyang, Banjar Regency.



Figure 2. Main sources of water pollution in Tamiyang Lakes a. floating net cages, b. rice fields, c. tourism area and floating latrines (Private documentation, 2023).

Tamiyang Lake is a lake that receives various waste directly or indirectly from the Riam Kanan Reservoir via the Riam Kanan River and runoff water from around the lake. The types of waste that enter Lake Tamiyang are thought to be similar to the waste that enters the Riam Kanan Reservoir, namely: a) resident waste; b) agricultural waste; c) industrial waste (gold) d) mining waste, e) livestock waste and f) floating net pond fishery activity waste [14]. The analysis results show that the biggest contribution to changes in water quality comes from floating net cages (71.2%), then gold mining (6.7%) and tourism areas (5%).



Picture 3. The source of pollution in Tamiyang Lake

Pollutant sources found in the Winongo River in Daerah Istimewa Yogyakarta consist of fish ponds, mixed gardens, domestic sewage canals, irrigation canals, plastic waste, latrines and chicken farms. Domestic waste ranks as the highest source of river pollution with a value of around 46% [9].

The Abundance of Phytoplankton

The phytoplankton diversity index in Tamiyang Lake is the high category with values ranging from 2.86 to 3.16 (Figure 4). The uniformity index is in the low category, with values ranging from 0.28 to 0.31. The dominance index ranges from 0.05 - 0.08. indicates that there is no dominant phytoplankton genus. The highest abundance of phytoplankton came from sand mining (31,524 Cells/L), and the lowest was from floating net cages (24,939 Cells/L).

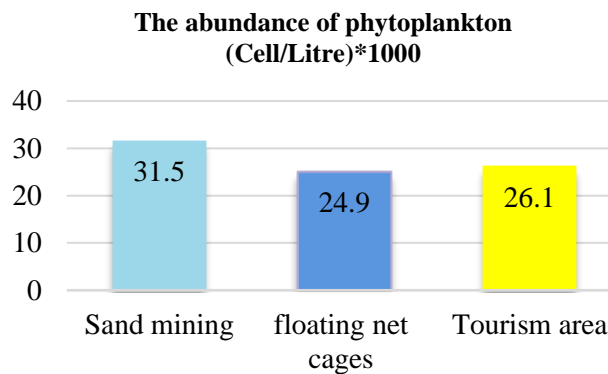


Figure 4. The abundance of pytoplankton in Tamiyang Lake of Banjar Regency

The ranges of phytoplankton abundance in the Perancak Estuary is 42,000 to 3,002,400 cells/m³. The highest abundance is found in the upstream and the lowest in downstream. The abundance of phytoplankton in aquatic can fluctuate in type and amount due to differences in nutrient content. Predation by zooplankton or herbivorous fish and tidal occurrence of sea water. Tides can affect plankton abundance and distribution. In high tide conditions. The flow of sea water to the river carrying the phytoplankton of sea water (Diatom) in it mixed with river water in which there is freshwater phytoplankton and vice versa. This is what causes Diatoms are found in the upstream which adds to the abundance of phytoplankton on the upstream and the discovery of Cyanophyceae at the stream and downstream. The phytoplankton abundance in Perancak Estuary varied between 33,600 – 3,002,400 cell/m³. Table 2 shows *Oscillatoria* sp. was dominance and distributes equally in all stations. However. the abundance of *Oscillatoria* sp. Has not reached endangers for the other lives organisms because according to [20] blooming will occurs when the abundance reaches 5x10⁵ to 2x10⁶

cells/l. Differences in the abundance of phytoplankton in Estuary Perancak are closely related to the physico-chemical conditions of the waters. According to [21]. Nutrients that get into the aquatic environment such as phosphate, nitrate, silica and ammonia will affect the development of phytoplankton and zooplankton [12],

The Environmental Parameters

There are six environmental parameters measured in this research, namely: Temp. (oC), pH, DO (mg/l), BOD (mg/l), Nitrates (mg/l) and Phosphates (mg/l) as shown in Table 1 below.

Table 1. Analysis of environmental parameters in Tamiyang Lake, Banjar Regency

Nu	Parameters	Site 1	Site 2	Site 3	PP No. 22 / 2021 Kelas 2
1	Temp. (°C)	29.6	30.1	29.5	Deviasi 3
2	pH	6.17	6.21	6.09	6 – 9
3	DO (mg/l)	2.21 *	2.67 *	3.4 *	4
4	BOD (mg/l)	44.14 *	14.21 *	66.26 *	3
5	Nitrates (mg/l)	0.01	0.01	0.012	10
6	Phosfates (mg/l)	0.18	0.15	0.12	0.2

Analysis of environmental parameters based on Chi-Square Test show that there is no difference between pH, temperature, DO, Nitrate, and Phosphate, except for BOD at a value of = 000 affected by the source of pollution. When compared, the results of Kudubun et al's (2020) research on the Winongo River, Yogyakarta complement the data obtained from this research. It is known that the parameters depth, brightness, TDS and TSS have a very strong correlation ($\text{sig} < 0.01$) with riparian vegetation. Meanwhile, the water discharge parameters, temperature, turbidity, DO, nitrate and phosphate have a strong correlation ($\text{sig} < 0.05$). The pH and DO parameters have a positive correlation with riparian vegetation, while the parameters of depth, water discharge, temperature, brightness, turbidity, TDS, TSS, nitrate and phosphate have a negative correlation with riparian vegetation [9]

Sustainable lake management is a commitment at the government level to save the lake from the environmental damage resulting from the exploration and exploitation of communities with an interest in lake waters [15]. Secara kualitatif tata guna lahan dan sumber pencemar memiliki pengaruh secara langsung terhadap kualitas air [9]

CONCLUSION

There is not a single source of pollution in the waters that dominates changes in water quality in Tamiyang Lake, Banjar Regency. The values of temperature, pH, nitrate and phosphate are still below the permissible water quality threshold.

REFERENCES

- [1] Bahri S, Kadir S, Suyanto and Lilimantik E 2019 Strategi Pengendalian Terhadap Polusi Air Sungai di SUB-DAS Riam Kiwa Kabupaten Banjar *EnviroScienteeae* **15** 291
- [2] Romanelli C, Cooper D, Campbell-Lendrum D, Maiero M, Karesh W B, Hunter D and Golden C D 2015 *Connecting Global Priorities : Biodiversity and Human Health A State of Knowledge Review* (World Health Organization and Secretariat of the Convention on Biological Diversity)
- [3] Bahri S, Ramadhan F and Reihannisa I 2015 Kualitas Perairan Situ Gintung, Tangerang Selatan *BIOGENESIS* **3** 16–22
- [4] Haryati S, Asmawi S, Yasmi Z and Bahan A 2019 Analisis Kualitas Air dan Tingkat Kesuburan Perairan Pada Kedalaman Berbeda Di Danau Tamiang Kecamatan Karang Intan Kabupaten Banjar Provinsi Kalimantan Selatan *Aquat. (Jurnal Tugas Akhir Mahasiswa)* **2** 116–25
- [5] Viani D Z and Retnaningdyah C 2018 Evaluasi Status Trofik dan Pencemaran Bahan Organik di Waduk Lahor Malang Menggunakan Bioindikator Diatom *Biotropika J. Trop. Biol.* **6** 10–5
- [6] Anas P, Jubaedah I and Sudinno D 2017 Kualitas Air dan Beban Limbah Karamba Jaring Apung di Waduk Jatiluhur Jawa Barat *J. Penyul. Perikan. dan Kelaut.* **11** 35–47
- [7] Sudinno D, Jubaedah I and Anas P 2015 Kualitas Air dan Komunitas Plankton Pada Tambak Pesisir Kabupaten Subang Jawa Barat *J. Penyul. Perikan. dan Kelaut.* **9** 13–28

-
- [8] Awaludin A S, Dewi N K and Ngabekti S 2015 Koefisien Saprobik Plankton di Perairan Embung Universitas Negeri Semarang *Indones. J. Math. Nat. Sci.* **38** 115–20
 - [9] Kudubun R, Kisworo and Rahardjo D 2020 Pengaruh Tata Guna Lahan, Tipe Vegetasi Riparian, dan Sumber Pencemar terhadap Kualitas Air Sungai Winongo di Daerah Istimewa Yogyakarta *Pros. Semin. Nas. Biol. di Era Pandemi COVID-19* 392–400
 - [10] Khaeriyah A and Burhanuddin 2015 Studi Kelimpahan dan Sebaran Phytoplankton Secara Vertikal di Pesisir Perairan Kuricaddi (untuk Peruntukan Budidaya Ikan dan Udang) *OCTOPUS J. Ilmu Perikan.* **4** 427–34
 - [11] APHA 2005 *Standard Methods for the Examination of Water and Wastewater* (Washington DC.: American Public Health Association/American Water Works Association / Water Environment Federation,)
 - [12] Hastuti A W, Pancawati Y and Surana I N 2018 The abundance and spatial distribution of plankton communities in Perancak Estuary, Bali *IOP Conf. Ser. Earth Environ. Sci.* **176**
 - [13] Pan J, Ma Y, Cai S, Chen Y and Chen Y 2022 Distribution patterns of lake-wetland cultural ecosystem services in highland *Environ. Dev.* **44** 100754
 - [14] Brahmana S S and Achmad F 2012 Potensi Beban Pencemaran Nitrogen, Fosfat, Kualitas Air, Status Trofik dan Stratifikasi Waduk Riam Kanan *J. Sumber Daya Air* **8** 53–66
 - [15] Prasetyo S, Anggoro S and Soeprbowati T R 2021 The Growth Rate of Water Hyacinth (*Eichhornia crassipes* (Mart.) Solms) in Rawapening Lake, Central Java *J. Ecol. Eng.* **22** 222–31

Variety of Ferns in Sultan Adam Grand Forest Park

Rina Falhiyah^{*(1)}, Aulia Ajizah⁽¹⁾, Amalia Rezeki⁽¹⁾

⁽¹⁾*Biology Education Faculty of Teacher Training and Education, Lambung Mangkurat University*

*Corresponding author: rinafalhiyah@gmail.com

Abstract: Sultan Adam forest park is located in the East Mandiangin Village area, Karang Intan District, Banjar Regency, South Kalimantan. Overall, this village is a humid hilly area, making a good habitat for ferns. This study aims to describe the types of ferns found in the Sultan Adam Mandiangin grand forest park area. The data were obtained using the total roaming method, then the morphological characteristics of each sample of ferns found were determined. Data analysis was carried out descriptively and identifying species with reference to the literature. The results of the study found 11 species of ferns, both epiphytes and terrestrial, belonging to 8 families, namely: *Pteris vittata* L., *Pityrogramma calomelanos*, *Lygodium circinnatum*, *Lygodium flexuosum*, *Selaginella ornata*, *Christella dentata*, *Nephrolepis exaltata*, *Pteridium aquilinum*, *Pyrrrosia adnascens*, *Drynaria sparsisora*, and *Asplenium nidus*. The results of this study are expected to provide information about the biodiversity of ferns in South Kalimantan.

Keywords: Variety of ferns, ferns, epiphytic ferns, terrestrial ferns, Sultan Adam Mandiangin Grand Forest Park

INTRODUCTION

Ferns are Cormophyta plants with spores that can live anywhere, the abundance and distribution of ferns is very high, especially in tropical rain forest areas. One of the tropical rainforest areas in South Kalimantan is the Sultan Adam Forest Park, Mandiangin. There are many typical tropical rainforest plants such as trees, shrubs, vines, epiphytic plants (plant species that live attached to other plants), mosses, microorganisms, algae, fungi, and ferns [18].

The South Kalimantan region generally has two seasons, namely the rainy season and the dry season. The rainy season ranges from October to May. The direction of the wind blows from the Northeast during the rainy season and every month has a wind speed of around 8-14 knots and an average of 5-6 knots per month. At air temperature, the maximum ranges from 33.1°C - 35°C, and the minimum ranges from 22.6°C - 23.8°C with an average temperature in the range of 15.6°C to 26.9°C. and has an average humidity of 77% - 91%. In South Kalimantan, the average solar radiation intensity is 52.5% with the highest increase in rainfall in March of 426.0 mm and the lowest decrease in rainfall of 75.0 mm in September [32]. According to the 2022 BPS Karang Intan, Sultan Adam Forest Park is located in the East Mandiangin Village area, Karang Intan District, Banjar Regency, South Kalimantan. Karang Intan District is located at coordinates 2°-3°49-43 '55-38" LS, 114°-115°E30-35'20-37". Overall, this village is a hilly area rich in flora and fauna.

During the rainy season, humidity is in the high range, so it really supports the growth of ferns. According to [11] ferns need sunlight and live in the open, are widely distributed, some live in groups, solitary and climb. Ferns form thickets that cover vacant lands, in closed areas with less light intensity and high air humidity. With such environmental conditions, the Sultan Adam forest park is a suitable habitat for ferns to live in, and based on the initial survey, various ferns were found. However, the question is what types of ferns are found in this forest park, and in which habitats can ferns samples be found.

This study aims to describe the types of ferns found in this forest park, and in the habitat samples of ferns can be found. The results of this study are expected to provide information about the biodiversity of ferns in South Kalimantan.

MATERIALS AND METHODS

Materials

The research used materials in the form of samples of ferns found on the ground (terrestrial ferns) and those attached to trees (epiphytic ferns) found in the Dutch Pesangrahan area in the Sultan Adam Mandiangin Grand Forest Park.

Methods

The research was conducted at the Dutch Pesanggrahan located in the Sultan Adam Mandiangin Forest Park from January to June 2023. Sampling of ferns was carried out using the total roaming method; namely by exploring the paths around the Dutch Pesanggrahan area with certain limitations such as steep geography and ravines. The samples observed included ferns found on the ground (terrestrial ferns) and those attached to trees (epiphytic ferns). When making observations, samples were photographed in their habitat, then taken and brought to the Laboratory to describe their morphological characteristics and to carry out microscopic observations and to identify the samples of ferns found. Identification of ferns refers to the books of Agatha et al (2019), Rismunandar and Ekowati, M (1991) and Sastrapradja, S (1979).

RESULTS AND DISCUSSION

Types of Ferns in Sultan Adam Mandiangin Forest Park

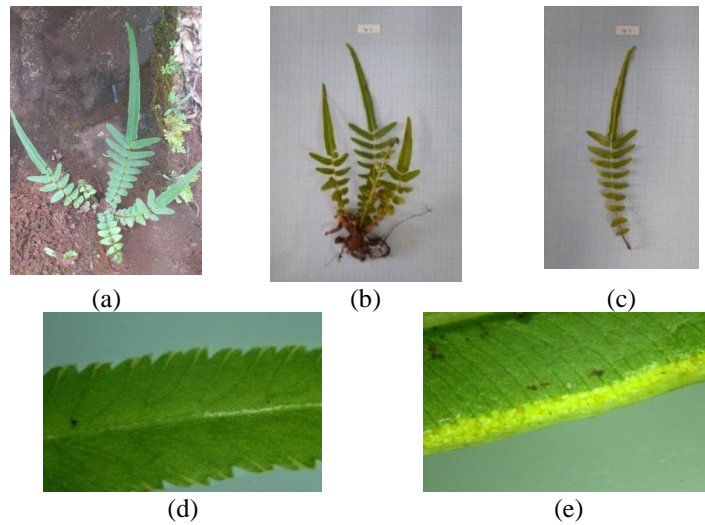
There are 11 species of ferns found in the Dutch Guest House of Sultan Adam Mandiangin Forest Park belonging to 8 families as shown in table 1. The habitats of ferns are varied, some live as epiphytes, live on the ground, on rocks, in water, live alone or in groups, to a large size that lives on rocks. Ferns have a fairly high adaptability, so it is not difficult to find ferns because they can live everywhere. Among them are in humid areas and under trees, on the banks of rivers, on steep slopes, in the mountains and many are even attached to tree trunks, rocks or growing on the ground [29].

Table 1. Types of ferns found

No.	Familia	Species	Habitat
1.	Pteridaceae	<i>Pteris vittata</i> L.	Terrestrial
2.	Pteridaceae	<i>Pityrogramma calomelanos</i>	Terrestrial
3.	Lygodiaceae	<i>Lygodium circinnatum</i>	Terrestrial
4.	Lygodiaceae	<i>Lygodium flexuosum</i>	Terrestrial
5.	Selaginellaceae	<i>Selaginella ornata</i>	Terrestrial
6.	Thelypteridaceae	<i>Christella dentata</i>	Terrestrial
7.	Dryopteridaceae	<i>Nephrolepis exaltata</i>	Terrestrial
8.	Dennstaedtiaceae	<i>Pteridium aquilinum</i>	Terrestrial
9.	Polypodiaceae	<i>Pyrrosia adnascens</i>	Epiphytes
10.	Polypodiaceae	<i>Drynaria sparsisora</i>	Epiphytes
11.	Aspleniaceae	<i>Asplenium nidus</i>	Epiphytes

Ferns are pioneer plants that are able to survive in various environmental conditions and grow in a variety of habitats both terrestrial, epiphytic and aquatic [26]. Ferns as part of biodiversity are plant communities that have important ecological functions in forest ecosystems, such as ground cover vegetation, litter mixtures for the formation of soil nutrients, and producers in the food chain, besides acting as a source of food and medicine. drugs [22]. The types of ferns that are commonly found are terrestrial. This is very possible because the Tahura area contains a lot of leaf litter on the soil surface, which means the availability of abundant nutrient sources due to the degradation of the litter. The ferns found in Sultan Adam Mandiangin Forest Park are as follows.

Pteris vittata L.



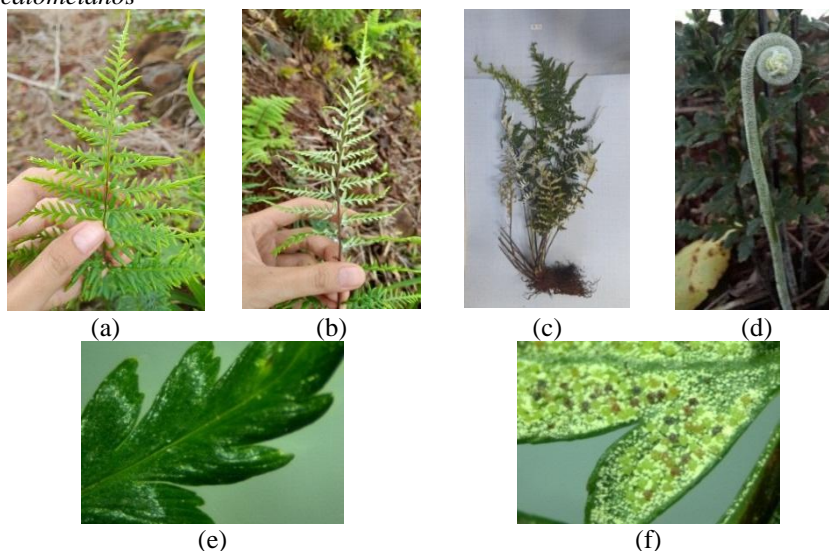
Gambar 1. *Pteris vittata* L. (a) habitus; (b) roots, stems, leaves; (c) leaf abaxial; (d) leaf edges; (e) sorus

Based on the results of observations, ferns *Pteris vittata* L has a terrestrial habitat which has fibrous roots with upright rhizome stem types. Leaf stalks are brownish green, round in shape with pinnate glans. Leaves are green with a rough surface. The leaf tips are blunt with a flat base, have a jagged leaf edge, and the layout of the leaves is opposite. The sorus is shaped like a line and is located on the edge of the underside of the leaf (abaxial).

According to [2] *Pteris vittata* has fibrous roots, upright rhizomes, and brown scales. Ental pinnate, the arrangement of leaves opposite or crossed, the tip of the leaf is pointed, the lowest leaf is the shortest and the longer the leaf goes up. According to [18], the leaves are green with a smooth surface, the petioles are brownish green and round in shape.

The sori are located along the leaf margins, but do not reach the tip or base [19]. Sorus is shaped like a line on the lower edge of the leaf. The habitat of *Pteris vittata* is terrestrial on flat ground, root gaps, gutter walls and sloping ground [2].

Pityrogramma calomelanos



Gambar 2. *Pityrogramma calomelanos* (a) dan (b) habitus; (c) roots, stems, leaves; (d) fiddleheads; (e) adaxial leaves (f) sorus

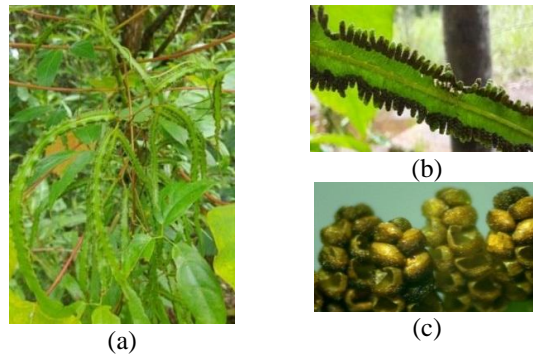
Based on observations, *Pityrogramma calomelanos* ferns have fibrous roots with upright rhizome stem types. Petiole blackish brown, round shape with bipinnate pinnatifid fronds. The leaves are light green,

the surface is smooth, the tip of the leaf is pointed with a tapered base, the leaf edges are wavy, and the leaf layout is alternate. Sori are round and located on the underside of the leaf (abaxial).

According to [18] *Pityrogramma calomelanos* has morphological characteristics of fibrous roots. The leaves are light green with a smooth leaf surface. The edges of the leaves are wavy with a pointed leaf tip and a tapered leaf base. The sori on these spikes are on the underside of the leaves and have a rounded shape (ovatus) [21].

According to [2] *Pityrogramma calomelanos* has upright rhizomes, brown scales. Ental bipinnate-pinnatifid, alternate leaf arrangement, fork venation, abaxial leaves covered with white scales. The stalks are green to brownish to shiny black, there are white scales. The habitat of this fern is terrestrial on river cliffs.

Lygodium circinnatum



Gambar 3. *Lygodium circinnatum* (a) habitus; (b) arrangement of sori (c) sorus

Based on observations, *Lygodium circinnatum* ferns have fibrous roots with creeping rhizome stem types. The leaf stalks are brownish green, round in shape with finger tips. The color of the leaves is green, the surface is smooth, the tip of the leaf is tapered with a rounded base, the leaf edge is flat, and the leaf layout is opposite. Sori are round and located on the under surface of the fertile leaves.

Lygodium circinnatum has fibrous roots [6]. Terrestrial habitat has a short rhizome that spreads on the ground. Stems are called rachis, up to 10 m long, 2-5 mm in diameter; rachis on sterile leaves branched like a fork, brownish green, direction of rotation to the right [23]. The stems are round and smooth and the direction of growth of the stems is climbing by using supports, namely other plants around it [34].

According to [30] *Lygodium circinnatum* can be distinguished from other types by looking at the form of the frond. Frond arranged pinnate, with the shape of fingers between 2-5, sometimes branching in two and each branching two more. The fronds are serrated and pale in color. In the fertile frond, the sporangia are located on the edges of the serrations of the frond.

Lygodium circinnatum leaves have a thin and slippery texture [6]. The base of the leaf is blunt, the flat leaf edge is somewhat wavy, the leaf tip is tapered with a finger shape, the upper surface of the leaf is dark green and the lower surface is light green [34]. Sori are brown to blackish in color and round in shape [39].

Lygodium flexuosum



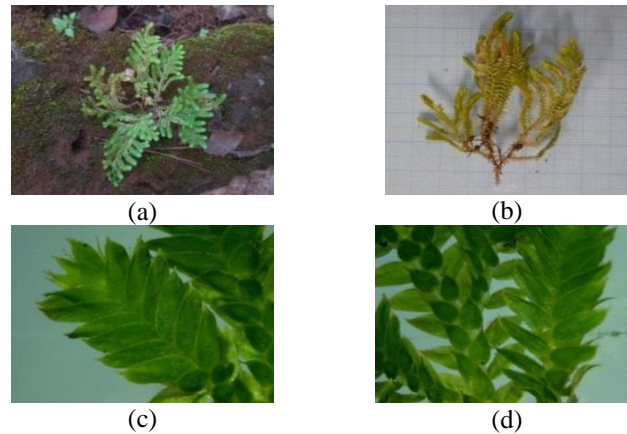
Gambar 4. *Lygodium flexuosum* (a) habitus; (b) leaf abaxial

Based on observations, *Lygodium flexuosum* ferns have fibrous roots with short creeping rhizomes. Petiole brown, round in shape with finger fronds. The leaves are dark green, the surface is smooth, the tip is tapered with a rounded base, the leaf edge is flat, and the leaf layout is finger-like.

Lygodium flexuosum has morphological characteristics of fibrous roots [37]. This fern has a short creeping rhizome with finger ridges [2]. Round stem shape with dark brown stem color [31]. Green leaves with a smooth leaf surface, flat leaf edges, alternating leaves, blunt tip with a tapered leaf base [37].

Sori are located along the leaf margins [31]. Sporangia are found on the marginal parts of the fertile leaves, are elongated, arranged in two rows on the leaf margins and are green in color [10]. The habitat of this fern is terrestrial on flat land and sloping land [2].

Selaginella ornata



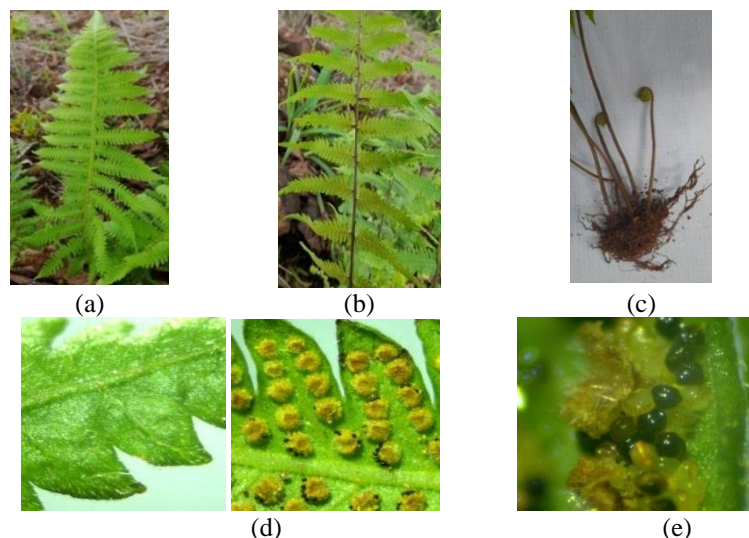
Gambar 5. *Selaginella ornata* (a) habitus; (b) roots, stems, leaves; (c) sporophyll leaves; (d) trophophyll leaves

Based on observations, the fern *Selaginella ornata* has a terrestrial habitat, has fibrous roots with creeping rhizome stems. Leaf stalks are green, round in shape with a single pinnate shape. The color of the leaves is green with a smooth surface. The tip of the leaf is tapered with a rounded base and has a flat leaf edge. The leaves are arranged alternately along the stem and the stem is bifurcated. At the time of observation, no strobilus was found.

According to [30] *Selaginella ornata* grows creepingly on the ground resembling moss. The leaves are arranged alternately along the stem. The leaves are fine textured. The trunk itself is usually forked and each branch forks again, and so on. The fertile leaves are arranged in a garland that resembles a grain called a strobilus. The strobili are located at the ends of the branches with a shape like a tassel with a whitish green color.

Selaginella ornata has a terrestrial habitat with single pinnate lobes and oval leaflets with dichotomous branching [11]. It has fibrous roots with spreading rhizomes. The leaf stalks are green, round in shape, the leaves are light green, the leaf surface is smooth, the leaf edge is flat with a pointed tip with a rounded base [18].

Christella dentata



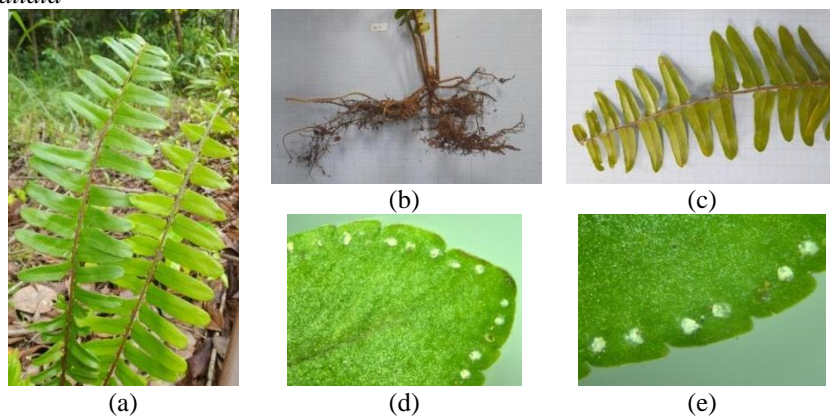
Gambar 6. *Christella dentata* (a) dan (b) habitus; (c) roots and fiddleheads; (d) upper and lower surfaces of leaves; (e) sorus

Based on observations, *Christella dentata* ferns have fibrous roots with upright rhizome stem types. Leaf stalks are green, round in shape with pinnate pinnatifid glans. The color of the leaves is green, the surface is smooth, the tip of the leaf is tapered with a notched base, the leaf edge is serrated, and the layout of the leaves is alternate. Sori are kidney-shaped and located on the underside of the leaf (abaxial).

According to [18] *Christella dentata* has fibrous roots with upright rhizomes. The color of the stalk is green with a round stalk shape. The color of the leaves is light green with a smooth leaf surface, serrated leaf edges, the location of the leaves is alternate, the tip of the leaf is tapered with a curved leaf base.

According to [2] *Christella dentata* has an erect and creeping rhizome, brown scales. Pinnate-pinnatifid enthals, alternating leaf arrangement, fork venation, slitted leaf edges, in sterile leaves the spacing between leaves is tight and in fertile leaves the spacing between leaves is stretched. The stalk is greenish brown, there are brown scales. Sori are kidney-shaped on the abaxial leaf, located on the superomedial side of the leaf, the sorus are arranged in two rows at each indentation of the leaf consisting of 2-4 sorus in each row. Habitat *Christella dentata* terrestrial on flat or sloping ground.

Nephrolepis exaltata



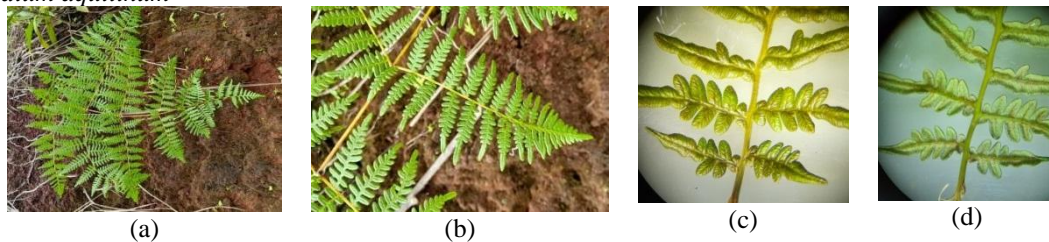
Gambar 7. *Nephrolepis exaltata* (a) habitus; (b) roots, stems and fiddleheads; (c) leaf abaxial; (d) and (e) sorus

Based on observations, *Nephrolepis exaltata* ferns have fibrous roots with upright rhizome stem types. Petiole brown, round in shape with pinnate globose. The color of the leaves is green, the surface is smooth, the tip of the leaf is blunt with a flat base, the edge of the leaf is wavy, and the layout of the leaves is alternate. Sori are round, white and located on the upper surface of the leaf (adaxial). *Nephrolepis* fern leaves have homophile properties, where the sporophyte is also a tropophyte [4].

Nephrolepis exaltata has a fibrous root, round in shape and has hairs on the root branches [3] with an upright rhizome stem type [36]. The color of the petiole is light brown, grows upright, round in shape, and has fine hairs on the young stalk. The color of the leaflets is green with a pointed tip, the base of the leaf is flat and has ears on one side, and the arrangement of the leaflets is alternate and opposite [3].

The leaf type of *Nephrolepis exaltata* is a single pinnate compound opposite (pinnate) with finely serrated leaf edges [24]. The tip of the leaf is blunt, the upper leaf surface is bright green while the underside is slightly faded in color with a soft texture [36]. Sorus is found on the upper surface of the leaves [3] which is round, irregularly distributed, white in color, and has no indusium [14]. Habitat *Nephrolepis exaltata* lives on cliffs and on rocky ground [24].

Pteridium aquilinum



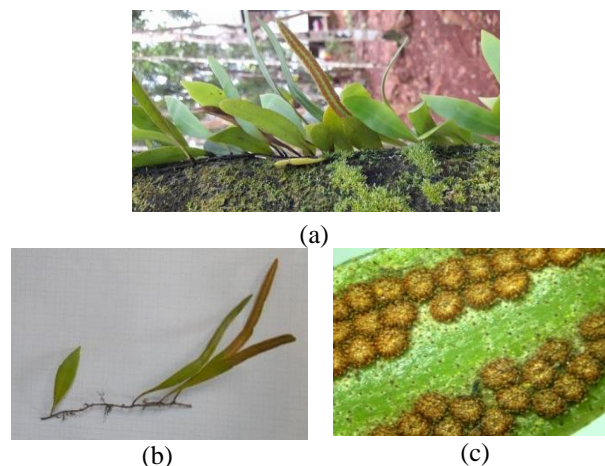
Gambar 8. *Pteridium aquilinum* (a) habitus; (b) leaf; (c) leaf adaxial (d) leaf abaxial

Based on observations, *Pteridium aquilinum* ferns have fibrous roots with a creeping rhizome stem type. Petiole brownish green, round shape with tri pinnate pinnatifid globular shape. The leaves are green, the surface is wrinkled, the tip of the leaf is pointed with a flat leaf base, the leaf edge is flat, and the layout of the leaves is alternate.

The fern plant *Pteridium aquilinum* has a terrestrial habitat. The stem is a rhizome, branched, spreads long underground [8] and has fibrous roots [13]. Petiole erect, round, light green. The leaves of *P. aquilinum* include odd pinnate compound leaves, light green in color, smooth or shiny leaf surface, elongated leaf shape with alternating arrangements, pointed leaf tips, toothed edges, and flat leaf bases [21].

P. aquilinum spores cover the stripes on the underside of the leaves. Sporangium aggregates into sori on the underside of the leaves [8]. The sori are protected by indusium, round in shape, located below the sinus between two indentations [27].

Pyrrosia adnascens



Gambar 9. *Pyrrosia adnascens* (a) habitus; (b) roots, stems, sterile leaves and fertile leaves; (c) sorus

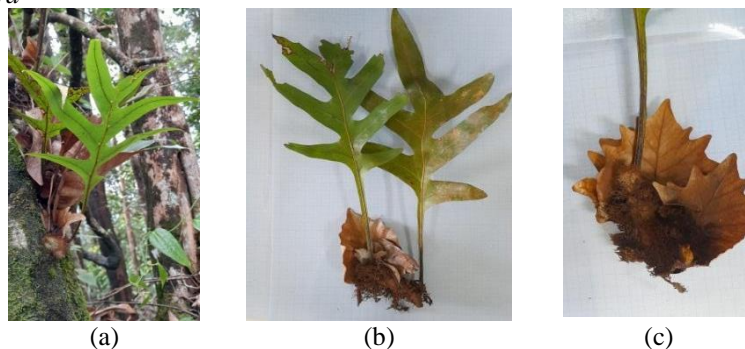
Based on observations, *Pyrrosia adnascens* ferns have epiphytic habitats, have fibrous roots with creeping rhizome stem types. Green leaf stalks, round shape with a single petal shape. The leaves are green, the surface is rough, the tip of the sterile leaf is pointed while the tip of the fertile leaf is rounded, the base of the leaf is pointed, the leaf edge is flat, and the layout of the leaves is alternate. Sori are round and located on the underside of the leaf (abaxial).

Pyrrosia adnascens ferns have a fibrous root system with a brown color, root length ranging from 3-5 cm. Long creeping rhizome stems, dark, brown scales that adhere tightly [1]. The leaf stalks are green and round in shape [38].

The leaves on this spike are single, lanceolate, arranged alternately, scaly, pale green in color, leaf surface texture is coarse, has flat edges, leaf base is rounded and leaf tips are pointed [1]. Shorter sterile leaves, oblong leaf shape with pointed leaf tips. Fertile leaves are longer, form a ribbon and have rounded leaf tips [16].

Sori are found along the area between the midrib and leaf margin. Sori are brown, round in shape and irregularly arranged [16]. The habitat for this fern is epiphyte [38].

Drynaria sparsisora



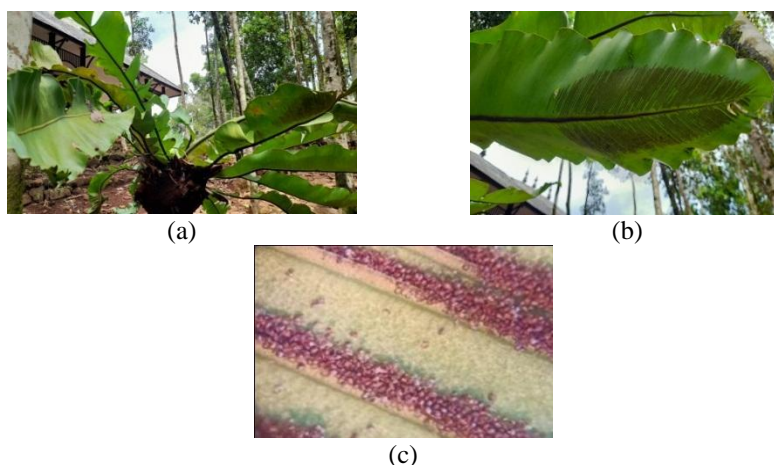
Gambar 10. *Drynaria sparsisora* (a) habitus; (b) roots, stalks, leaves; (c) nest leaves

Based on the results of observations, *Drynaria sparsisora* ferns have epiphytic habitats, have fibrous roots with creeping rhizome stem types. Petiole brown, semicircular with pinnatifid frond. The leaves are green, the surface is green, the tip and base of the leaf is pointed, the leaf edge is flat, and the leaf layout is alternate.

Drynaria sparsisora ferns have fibrous roots [7]. Long creeping rhizomes, dense scaly, shield-shaped scales, light brown to dark. Ental pinnatifid, wavy edge pointed tip, venation mesh. Stem reddish brown to dark, tough, winged [2]. This fern stalk has a semicircular shape [7].

Drynaria sparsisora leaves are light green to dark green with a smooth leaf surface. Flat leaf edges with alternating leaf arrangements with leaf tips and leaf bases that are pointed. These ferns live attached to their host like trees [7]. Finatifid nest leaves, young brown, stiff, tough sit covering the rhizomes. Sorus round, brown without indusium, on an abaxial surface from tip to center of leaf [2].

Asplenium nidus



Gambar 11. *Asplenium nidus* (a) habitus; (b) underside of leaves; (c) sori

Based on observations, *Asplenium nidus* ferns have fibrous roots with upright rhizome stem types. Petiole blackish brown, round shape with a single petal shape. The leaves are green, the surface is smooth, the leaf tip is blunt, the leaf base is pointed, the leaf edge is wavy, and the leaf arrangement is rosette. The sori are striped, brown in color and are located on the abaxial surface of the leaf and near the midrib.

Asplenium nidus has morphological characteristics, namely fibrous roots. It has dark brown leaf stalks and is round in shape [18]. It has an upright and short rhizome and has a single enthal [2].

The leaves are arranged in a rosette with a blunt tip and a pointed base [18]. Sori are formed on the underside of the leaves, scattered in lines, or elongated but narrow [4]. The sori are brown in color with the shape of lines arranged extending from the center of the leaf vein to the edge [17] which is located below the leaf surface and is in the leaf veins [20]. *Asplenium nidus* ferns have epiphytic and terrestrial habitats on flat land and are planted in pots [2].

Supporting Environmental Factors

The survival of ferns is influenced by the environmental conditions of the habitat in accordance with the adaptability of these plants. The state of the habitat environment can be seen from the results of environmental parameter measurements. The environmental parameters measured were light intensity, wind speed, soil pH, soil moisture, air humidity, altitude and air temperature. The results of environmental parameter measurements can be seen in table 2.

Table 2. Table of environmental parameters

No	Parameter	Tools	Unit	Range			Literature	
				1	2	3		
1.	Light Intensity	4 in 1 luxmeter	Lux	Min	8890	8670	8450	1037-8607(*)
				Max	14,590	18,540	15,960	
2.	Wind Speed	4 in 1 Anemometer	m/s	Min	0.0	0.0	0.0	0,1 – 1,4 (**)
				Max	0.0	0.0	0.0	
3.	Soil pH	Soil tester	-	6.4	6.5	6.5	5.5 - 6.5 (***)	
4.	Soil moistur		%	76	80	78	5 - 80(***)	
5.	Air Humidity	Hygrometer	%	73	77	71	60 - 80 (****)	
6.	Place Altitude	Altimeter	mdpl	315	315	315	150-250 (*****)	
7.	Air Temperature	Thermometer	°C	27	29	28	25 - 35 (***)	

(*)[18]

(**)[9]

(***)[28]

(****)[15]

(*****)[33]

Based on the results of measurements taken during the day, light intensity is detected in the range of 8450-8890 lux. According to [18] the light intensity of fern growth in the morning is 1037 lux, during the day it is 8607 lux, and in the evening it is 5650 lux. According to [15] ferns grow well in shaded conditions. Good light intensity for ferns growth ranges from 200-600 fc.

Based on the measurement results, the wind speed is 0.0 m/s. According to [9] the wind speed that is good for the growth of ferns has a range of 0.1-1.4 m/s. According to [12] wind speed greatly affects the existence of ferns. This is because the spores are spread by the wind. The greater the wind speed, the farther the spores will fly, and vice versa if the wind speed is low, the spores will only fly around it.

Soil pH measurements at the study site ranged from 6.4 to 6.5. According to [28] ferns are suitable for growing with soil pH ranging from 5.5-6.5. Soil acidity is influenced by several factors such as the content of elements of copper (Cu), aluminum (Al), and iron (Fe). In addition, soil acidity is also caused by the decomposition of organic matter which releases calcium (Ca). High rainfall can also affect soil acidity due to leaching of soil nutrients.

Soil moisture measurements at the study site ranged from 76-80%. According to [28] the optimum soil moisture for the growth of ferns is 50% -80%. According to [12] ferns prefer shady places with a high degree of humidity, so many types of ferns were found at the research site.

Air humidity measurements at the study site ranged from 71-77%. According to [15] the optimal humidity for the growth of ferns ranges from 60% -80%. Meanwhile, the lowest humidity that can be tolerated by ferns is around 30%. According to [28] tropical forests have a humid environment so they are very suitable for plants to help the photosynthetic cycle needed for development, especially for ferns.

Based on the measurement results, the altitude is 315 meters above sea level. According to [33] an altitude of 150-250 meters above sea level is the most suitable growing place for terrestrial ferns and epiphytic ferns, this is due to the elevation above sea level, high humidity and suitable temperature. There are more ferns in the highlands than in the lowlands, this is due to the presence of higher humidity, more water flow, the presence of fog, and even the amount of rainfall affects the species. According to [5], an altitude of more than 2000 m asl can result in a reduction in the number of fern species found. This is due to the reduced number of trees as shade, resulting in higher intensity of the sun and wind, so that only certain ferns can survive.

Based on the measurement results, the air temperature is in the range of 27-29°C. According to [15] the optimal air temperature for the growth of ferns is 25-35°C. According to [25] the temperature in tropical rain forests is relatively low because the speed of the wind that blows is also low so that water vapor from

plant evaporation and oxygen produced by plants are not blown away by the wind. At this temperature it is very supportive of the survival of ferns.

CONCLUSION

There are 11 species of ferns found in the Dutch Pesangrahan in the Sultan Adam Mandiangin Forest Park belonging to 8 families, both epiphytic and terrestrial, namely, *Pteris vittata* L., *Pityrogramma calomelanos*, *Lygodium circinnatum*, *Lygodium flexuosum*, *Selaginella ornata*, *Christella dentata*, *Nephrolepis exaltata*, *Pteridium aquilinum*, *Pyrrosia adnascens*, *Drynaria sparsisora*, and *Asplenium nidus*.

ACKNOWLEDGEMENT

A big thank you goes to the Manager of the Sultan Adam Mandiangin Grand Forest Park who has given permission to collect research data. Thank you also to all the research team friends who have helped, provided support in carrying out the research, and all those who have helped.

REFERENCES

- [1] G.S. Adjie., A. Ajizah dan S. Amintarti, *Pteridophyta di Sekitar Danau Sari Embun Kabupaten Tanah Laut dan Pengembangan Buku Saku* (JUPEIS: J. Pendidikan dan Ilmu Sosial, 2022), 1(3), pp. 230-240.
- [2] S. M. Agatha., K. A. Safitri., A. Pulungan., Maskana., dan A. Sedaya, *Panduan Lapangan Paku-Pakuan (Pteridophyta) Di Taman Margasatwa Ragunan* (Laboratorium Biologi Fmipa Universitas Negeri Jakarta, Jakarta Timur, 2019).
- [3] S. Q. Aini., S. Ifadatin dan Z. Zakiah, Karakteristik Morfologi Pada Tumbuhan Paku *Nephrolepis biserrata* (Sw.) Schott dan *Nephrolepis exaltata* (L.) Schott Di Kawasan Kampus Universitas Tanjungpura (*Protobiont*, 2022), 11 (1), pp. 11-16.
- [4] A. Ajizah dan C. Pratama, *Tumbuhan Paku di Area Kampus ULM Banjarmasin* (Penerbit Lambung Mangkurat University Press, Banjarmasin, 2020).
- [5] J. S. Anwar., N. Damanik., Hisyam dan A.J. Whitten, *Ekologi Ekosistem Sumatera* (UGM Press, Yogyakarta, 1987).
- [6] J. Betty., R. Linda dan I. Lovadi, *Inventarisasi Jenis Paku-Pakuan (Pteridophyta) Terrestrial Di Hutan Dusun Tauk Kecamatan Air Besar Kabupaten Landak* (*Jurnal Protobiont*, 2015), 4(1), pp. 94-102.
- [7] N. Dina., S. Amintarti dan A. P. Putra, *Pengembangan Modul Elektronik Materi Pteridophyta Di Taman Buah Lokal Kawasan Mangrove Rambai Center Sebagai Bahan Ajar Di SMA* (JUPEIS: Jurnal Pendidikan dan Ilmu Sosial, 2022), 1(3), pp. 173-187.
- [8] D. F. Hanas., E. J. Bria dan I. M. Y. Tnunay, *Inventarisasi Tumbuhan Paku (Pteridophyta) di Dusun Oelmuke, Desa Tasinifu, Kecamatan Mutis Kabupaten Timor Tengah Utara, Provinsi Nusa Tenggara Timur* (*Jurnal Saintek Lahan Kering*, 2019), 2(2), pp. 30-32.
- [9] M. Hanif, "Analisis Vegetasi Tumbuhan Paku (Pteridophyta) Di Taman Kota 2 BSD Kota Tangerang Selatan," Skripsi, Universitas Islam Negeri Syarif Hidayatullah, Jakarta, 2022.
- [10] H. Hasibuan., Rizalinda dan R. P. W. Elvi, *Inventarisasi jenis paku-pakuan (Pteridophyta) di hutan sebelah darat Kecamatan Sungai Ambawang Kalimantan Barat* (*J. Protobiont*, 2016), 5(1), pp. 46-58.
- [11] M. A. Hutasuhut dan F. Husnarika, *Keanekaragaman Paku-pakuan Terrestrial di Kawasan Taman Wisata Alam Sicike-Cike* (*Jurnal Biolokus: Jurnal Penelitian Pendidikan Biologi dan Biologi*, 2019), 2(1), pp. 146-157.
- [12] R. Imaniar., Pujiastusti dan S. Murdiah, *Paku Di Kawasan Air Terjun Kapas Biru Kecamatan Pronojiwo Kabupaten Lumajang Serta Pemanfaatannya Sebagai Booklet* (*Jurnal Pendidikan Biologi*, 2017), 6 (3), pp. 337-345.
- [13] P. Imat., A. G. Maulidiah dan A. Liana, Identifikasi Tumbuhan Paku di Situs Wisata Air Terjun Bantimurung (CELEBES BIODIVERSITAS: J. Sains dan Pendidikan Biologi, 2020), 3(1), pp. 35-39.
- [14] M. Jannah dan N. Sofiyanti, Karakterisasi Morfologi 3 Jenis Paku *Nephrolepis* (*Nephrolepidaceae*) Di Perkebunan Kelapa Sawit PT. Panca Surya Garden Di Kabupaten Kampar Provinsi Riau, 2020.
- [15] N. Komaria, *Identifikasi dan Inventarisasi Tumbuhan Paku Epifit di Lingkungan Kampus Universitas Jember Untuk Penyusunan Buku Nonteks* (Universitas Jember, Jakarta, 2015).
- [16] W. F. Lindsari., R. Linda dan I. Lovadi, *Jenis-Jenis Paku Epifit di Hutan Desa Beginjan Kecamatan Tayan Hilir Kabupaten Sanggau* (*Protobiont*, 2015), 4 (3), pp. 65-73.

- [17] R. Listiyanti., S. Indriyani dan N. Ilmiah, *Karakteristik Morfologi Jenis-Jenis Paku Epifit Pada Tanaman Kelapa Sawit Di Desa Tegalrejo* (Al Kawnu: Science And Local Wisdom Journal, 2022), 2 (1), pp. 99-106.
- [18] A. Majid., A. Ajizah dan S. Amintarti, *Keragaman Tumbuhan Paku (Pteridophyta) Di Taman Biodiversitas Hutan Hujan Tropis Mandiangin* (Jurnal Al-Azhar Indonesia Seri Sains Dan Teknologi, 2022), 7(2), pp. 102-12.
- [19] M. Mumpuni, *Variasi Morfologi Pteris vittata l. (Pteridaceae; Pteridophyta) dan Korelasinya dengan Ketinggian Lokasi Tempat Tumbuhnya di Jawa* (BIOLINK: Jurnal Biologi Lingkungan Industri Kesehatan, 2016), 2(2), pp. 100-109.
- [20] L. Nafili dan T. M. Sarjani, *Identifikasi Letak Dan Bentuk Sorus Pada Tanaman Paku (Pteridophyta) di Taman Hutan Raya Bukit Barisan Kecamatan Dolatrayat Kabupaten Karo* (Jurnal Jeumpa, 2019), 6(2), pp. 226-235.
- [21] N. Nurchayati, *Identifikasi Profil Karakteristik Morfologi Spora dan Protalium Tumbuhan Paku Famili Polypodiaceae* (Bioedukasi, 2014), 14(2).
- [22] A. Pradipta., R. Saputri., S. D. Ami dan A. Walid, *Inventarisasi Jenis Tumbuhan Paku (Pteridophyta) di Desa Padang Pelasan Kabupaten Seluma* (Jurnal Biosilampari, 2020), 3(1), pp. 13-19.
- [23] M. Rahayu., E. S. Kuncari dan M. Setiawan, *Ethnobotanical Study of Lygodium circinnatum and its Utilization in Crafts Weaving in Indonesia* (Biodiversitas Journal of Biological Diversity, 2020), 21(2), pp. 617-621.
- [24] R. Ramadani., Z. I. Navia dan A. Y. Persada, *Inventarisasi Jenis Paku Terrestrial (Pteridophyta) Di Kawasan Wisata Air Terjun Tujuh Tingkat Desa Selamat Kecamatan Tenggulun Kabupaten Aceh Tamiang, Aceh* (Al-Hayat: Journal of Biology and Applied Biology, 2021), 4(2).
- [25] D. S. Ridianingsih., P. Pujiastuti dan S. A. Hariani, *Inventarisasi Tumbuhan Paku (Pteridophyta) Di Pos Rowobendo-Ngagelan Taman Nasional Alas Purwo Kabupaten Banyuwangi* (Bioeksperimen: Jurnal Penelitian Biologi, 2017), 3(2), pp. 20-30.
- [26] Rismunandar dan M. Ekowati, *Tanaman Hias Paku-pakuan* (Penebar Swadaya, Jakarta, 1991).
- [27] Rugayah., S. Sunarti., D. Sulistiarini., A. Hidayat dan M. Rahayu, *Daftar Jenis Tumbuhan di Pulau Wawonii, Sulawesi Tenggara* (LIPI Press, Jakarta, 2015).
- [28] R. Rukmana, *Perawatan dan Pembibitan Paku Hias* (Kanibus, Yogyakarta, 1997).
- [29] D. E. Sahertian dan L. D. Tetelepta, *Inventarisasi Tumbuhan Paku (Pteridophyta) di Kawasan Benteng Duurstede Desa Saparua Kabupaten Maluku Tengah* (Jurnal Ilmu Alam dan Lingkungan, 2022), 13(1).
- [30] S. Sastrapradja., J. J. Afriastini., D. Darnaedi dan E. A. Widjaya, *Jenis Paku Indonesia* (Lembaga Biologi Nasional-LIPI, Bogor, 1979).
- [31] H. M. Saswita dan N. Sofiyanti, *Karakterisasi Morfologi 3 Jenis Paku Dari Genus Lygodium (Lygodiaceae) Di Kecamatan Bunga Raya, Kabupaten Siak, Provinsi Riau*, 2019.
- [32] A. S. E. Sudrajat dan S. Subekti, *Pengelolaan ekosistem gambut sebagai upaya mitigasi perubahan iklim di Provinsi Kalimantan Selatan* (Jurnal Planologi, 2019), 16(2), pp. 219-237.
- [33] S. Surfiana., S. Kamal dan M. Hidayat, *Keanekaragaman Tumbuhan Paku (Pteridophyta) Berdasarkan Ketinggian Di Kawasan Ekosistem Danau Aneuk Laot Kota Sabang* (Prosiding Biotik, 2018), 5(1), 452-459.
- [34] E. Taslim., R. Ramadanil dan S. M. Sulaeman, *Inventarisasi Jenis Paku-pakuan (Pteridophyta) Terrestrial di Jalur Pendakian Nokilalaki Kawasan Taman Nasional Lore Lindu* (Biocelbes, 2019), 13(2).
- [35] G. Tjitrosoepomo, *Taksonomi Tumbuhan Schizophyta, Thallophyta, Bryophyta, Pteridophyta* (Gadjah Mada University Press, Yogyakarta, 2014).
- [36] M. S. Yusal dan G. Toni, *Fern inventorization in Cunca Rami Waterfall Zone of West Manggarai, East Nusa Tenggara* (Jurnal Pembelajaran Dan Biologi Nukleus, 2021), 7(1), pp. 218-234.
- [37] A. Majid, "Pengembangan Modul Pteridophyta Berbasis Keragaman Tumbuhan Paku Di Taman Biodiversitas Hutan Hujan Tropis Mandiangin Sebagai Bahan Ajar Di SMA," Skripsi, Universitas Lambung Mangkurat, Banjarmasin, 2021.
- [38] G. S. Adjie, "Pteridophyta Di Sekitar Danau Sari Embun Kabupaten Tanah Laut Dan Pengembangan Buku," Skripsi, Universitas Lambung Mangkurat, Banjarmasin, 2022.
- [39] N. Khoirunnisa dan Nurmiyati, *Karakteristik Sorus Pteridophyta Di KHDTK Gunung Bromo Kabupaten Karanganyar* (In Proceeding Biology Education Conference, 2022), 19 (1), pp. 14-22.

Toxicity of Methanol Extract from Kitolod (*Hippobroma Longiflora* (L.) G. Don) Leaves and Stems to Mortality Rate of Cocoa Mealybug (*Planococcus Minor* Maskell., Hemiptera: Pseudococcidae)

Rochmah Agustina^{*(1)}, Nismah Nukmal⁽¹⁾, C.N. Ekowati⁽¹⁾, and Gina Dania Pratami⁽¹⁾

⁽¹⁾Biology Department, FMIPA, Universitas Lampung
Jln. Soemantri Brojonegoro No 1 Gedong meneng, Kec. Rajabasa,
Kota Bandar Lampung, Lampung 35141

*Corresponding author: rochmah.agustrina@fmipa.unila.ac.id

Abstract. Cocoa mealybug (*Planococcus minor* Maskell.) is one of cocoa plants pest in Indonesia. Kitolod (*Hippobroma longiflora* (L.) G. Don) is known to contain secondary metabolites that have bioactivity as insecticides. This study aims to determine the methanol extracts toxicity of kitolod leaves and stems on cocoa mealybug mortality. Toxicity tests of kitolod leaf and stem extracts were conducted separately, using a 2-factor randomized group design (RGD) with 5 replicates. Each replicate used 10 mealybugs. First factor, kitolod extract concentration, consisting of 1%, 2%, 3%, 4%, and positive control (25% methomyl). The second factor, observation time consisted of 12, 24, 48, and 72 hours after treatment. The data obtained were analyzed by probit EXE to determine the LC₅₀ and LT₅₀ values, ANOVA using SPSS 25 application, followed by Tukey test at $\alpha = 0.05$. The results showed that LC₅₀ for kitolod leaf extract was 3.065%, but LC₅₀ was not obtained from stem extract. The LT₅₀ for kitolod leaf and stem extracts was at a concentration of 3%, with a faster time of mealybug death by the leaf extract than by the stem extract, which was 2.371 and 7.024 hours, respectively. Thus, the kitolod leaf extract has higher toxicity than the stem extract.

Keywords:

INTRODUCTION

Indonesia is one of the third largest cocoa producing countries in the world, and cocoa beans are an important plantation commodity in the world market (1). Cocoa beans are the main ingredient of food and beauty industry products that have a high antioxidant content so that they have a high economic value (2). From 2018 to 2020 the yield of Indonesian cocoa plantations showed a decline (3).

The decline in cocoa production value is thought to be caused by plant disrupting organisms (pests). One of the pests that attack cocoa plants is the cocoa mealybug (*Planococcus minor* Maskell.). Mealybugs will attack the base of the fruit and suck phloem fluid from the small fruit, thus inhibiting further fruit growth and causing the fruit to dry out (4). Efforts to control mealybug pests are generally carried out by using chemical insecticides. However, the use of this type of insecticide often leaves residues that can pollute the environment. A strategy to reduce the negative impact of using chemical insecticides is to utilize bio-insecticides or natural insecticides obtained from plants (5). According to Fitri (5) bioactive compounds contained in plants that are toxic are known to be effective in killing pests so that they can be used as bio-insecticides are alkaloids, flavonoids, and saponins.

According to Ramadhona (7), alkaloid, saponin, flavonoid and terpenoid compounds in papaya leaves have the potential to control aphids that attack eggplants during the vegetative phase, because these compounds are known to be able to damage nerve tissue and affect several physiological systems of aphid development. Meanwhile, Nukmal and Andriyani (8) proved that flavonoid compounds from the coumarin group in gamal leaves (*Gliricidia sepium*) are effective as bio-insecticides on cocoa mealybugs because they can irritate the skin and inhibit the transportation of amino acid leucine.

Kitolod (*Hippobroma longiflora* (L.) G. Don) is known to contain bioactive compounds such as alkaloids, flavonoids, and saponins which are secondary metabolic products (9). In this study, the toxicity test of methanol extract of kitolod leaves and stems against cacao mealybugs was carried out to determine its potential as a bio-insecticide to control insect pests in plants, specifically in this study is to control cocoa mealybugs.

MATERIALS AND METHODS

Materials

Kitolod plants (*Hippobroma longiflora* (L.) G. Don) were collected from areas around residential areas in Bandar Lampung at coordinates 5°20' - 5°30' South latitude and 105°28'-105°37' East longitude and in Pesawaran at coordinates 104°92' - 105°34' East longitude, and 5.12o - 5.84o South latitude. Cocoa mealybugs (*Planococcus minor* Maskell.) and cocoa beans were obtained from cocoa plants growing around Pesawaran. Chemical insecticides methomyl 25% and methanol were obtained from agricultural and chemical stores in Bandar Lampung, respectively. The equipment used are various glassware and other laboratory equipment used to extract bioactive compounds from kitolod plants and toxicity tests.

Toxicity tests of methanol extracts of kitolod leaves and stems were carried out separately and each was arranged in a 2-factor factorial experimental using a Randomized Block Design (RBD). The first factor was methanol extract of kitolod with concentrations of 1%, 2%, 3%, 4%, and positive control (25% methomyl, chemical insecticide). Determination of extract concentration follows Prijono (2005) who argues that the concentration of methanol dry extract is determined based on the standard standard of effective extract which is $\leq 5\%$. The second factor was observation time of mealybug mortality consisting of 12, 24, 48, and 72 hours after treatment. The number of mealybugs used for each replication was 10 (10).

Kitolod extraction

The preparation of kitolod extract followed the method in Nukmal and Andriyani (8) and was carried out at the Botany Laboratory of the Biology Department and the Organic Chemistry Laboratory of the Chemistry Department of the Faculty of Mathematics and Natural Sciences, University of Lampung. Fresh kitolod plants that have been cleaned from soil and other impurities are dried for 7-10 days without sunlight. After drying, the leaves and stems of kitolod were ground separately into powder which was then extracted separately. 500 grams of kitolod powder was soaked in 5 liters of methanol in a tightly closed beaker glass for 3x24 hours to obtain methanol macerate. The methanol macerate obtained was evaporated using a rotary vacuum evaporator at a temperature of 40°C with a pressure of 122 mbar until the methanol extract was obtained in the form of a paste and then weighed to determine its weight.

Preparation of kitolod extract concentration

The concentration of kitolod extract to be tested was made by diluting the kitolod methanol extract paste with distilled water. Determination of 1%, 2%, 3% and 4% (w/v) concentration of kitolod extract followed the method used by Rusman et al. (2020) as follows:

$$\% = w/v$$

% = percent of substance (%)

w = mass of solute (gr)

v = solvent volume (ml)

Before dissolving with distilled water, the paste that has been weighed as needed is first dripped with tween 20 solution little by little until it is completely dissolved, then diluted with distilled water until it reaches the volume corresponding to the predetermined extract concentration.

Toxicity test of kitolod methanol extract

Mealybugs that will be used for toxicity tests are acclimatized by placing whole cocoa pods infested with mealybugs in 2000 ml beaker glass covered with gauze for 1 day in the laboratory (12). The insecticidal activity test against mealybugs was conducted following the residual effect method (8). Pieces of exocarp and mesocarp of cocoa pods weighing about 30 grams were soaked for 10 minutes in kitolod extract with the appropriate concentration. Then 10 female mealybugs that had been acclimatized for 1 day before treatment were placed on the pieces of cocoa pods and maintained in the test container. Observations of mealybug mortality were made at 12, 24, 48, and 72 hours after treatment (10).

RESULTS AND DISCUSSION

The toxicity of methanol extract of leaves and stems of kitolod on the mortality rate of cocoa mealybugs was observed based on LC₅₀ and LT₅₀ of the extract of leaves and stems of kitolod. Lethal Concentration 50 (LC50) is the concentration that causes death in 50% of experimental animals while Lethal Time 50 (LT50) is the time required to kill 50% of the test animal population. In this study the test animal used was mealybug.

The anova result shows that both methanol extract of kitolod leaves (Table 1a) and methanol extract of kitolod stem (Table 1b) have significant effect on mealybug mortality for extract concentration (C), observation time (OT), and the interaction between concentration and observation time (C x OT).

Table 1. Anova result of methanol extract of kitolod leaves (a) and stems (b) on mealybug mortality

a. Leaf extract of kitolod

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	359.250 ^a	19	18.908	27.670	.000
Intercept	1000.417	1	1000.417	1464.024	.000
Concentration (C)	91.833	4	22.958	33.598	.000
Observation time (OT)	241.250	3	80.417	117.683	.000
C * OT	26.167	12	2.181	3.191	.003
Error	27.333	40	.683		
Total	1387.000	60			
Corrected Total	386.583	59			

R Squared = .929 (Adjusted R Squared = .896)

b. Stem extract of kitolod

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	337.250 ^a	19	17.750	39.444	.000
Intercept	183.750	1	183.750	408.333	.000
Concentration (C)	286.167	4	71.542	158.981	.000
Observation time (OT)	29.650	3	9.883	21.963	.000
C * OT	21.433	12	1.786	3.969	.000
Error	18.000	40	.450		
Total	539.000	60			
Corrected Total	355.250	59			

R Squared = .949 (Adjusted R Squared = .925)

Table 2 shows that the highest number of mealybug mortality due to concentration treatment was obtained from the positive control treatment (25% methomyl) which was significantly higher than the concentration treatment of kitolod methanol extract. The results of the concentration treatment of methanol extracts of kitolod from both leaves and stems showed that up to a concentration of 3% kitolod extract, the number of mealybug mortality increased as the concentration of the extract tested increased. However, at a concentration of 4%, both methanol extracts of kitolod leaves and stems, the number of mealybug mortality decreased, and the number was significantly lower than the number of mealybug mortality due to the 3% concentration treatment. Based on the average number of mortality, overall, it can be seen that in all kitolod extract concentration treatments, the number of mealybug mortality caused by methanol extract of kitolod leaves was higher than that caused by methanol extract of kitolod stems.

Table 2. Mean of mealybug mortality resulted from methanol extract of kitolod leaves and stems

Concentration of Extract methanol	Mealybug mortality ($\bar{X} \pm Sd$)	
	of kitolod leaves	of kitolod stems
1%	2.67 ± 1,37 ^a	0.58 ± 0,52 ^{ab}
2%	3.92 ± 1,98 ^{bc}	0.58 ± 0,5 ^{ab}
3%	4.75 ± 2,83 ^c	1.17 ± 1,12 ^b
4%	3,00 ± 1,91 ^{ab}	0.33 ± 0,49 ^a
Positive control	6.08 ± 3,06 ^d	6.08 ± 2,07 ^c

Note: Mean values followed by the same uppercase letter in the same column indicate not significantly different at the $\alpha = 5\%$ level.

Table 3 shows the average number of mealybug mortalities at different observation times. Mealybug mortality increased with increasing observation time where the highest mortality was obtained at 72 hours observation time. As in the data on the number of mortalities due to the treatment of methanol extract concentration of kitolod, from the data

in Table 3, it can also be seen that for each observation time, the number of mealybug mortality caused by leaf extract treatment is higher than that caused by kitolod stem extract.

Table 3. Average mortality of cocoa mealybugs at different observation times

Observation time	Mealybug mortality ($\bar{X} \pm Sd$)	
	kitolod leaves	kitolod stems
12 hours after treatment	1.93± 0.89 ^a	0.67± 1.45 ^a
24 hours after treatment	2.40± 0.91 ^a	1.60± 2.16 ^b
48 hours after treatment	5.20± 1.82 ^b	2.27± 2.93 ^c
72 hours after treatment	6.80± 2.34 ^c	2.47± 2.80 ^c

Note: Mean values followed by the same uppercase letter in the same column indicate not significantly different at the $\alpha = 5\%$ level.

The interaction treatment between concentration and observation time (C x OT) in Table 2 shows that the observation time of 72 hours after the positive control (25% methomyl) causes the highest mealybug mortality for both methanol extracts of kitolod leaves and stems (Table 4).

Table 4. Average mortality of cacao mealybugs due to the interaction of concentration treatment and observation time.

Observation time	Source of Extract	Mean of mealybug mortality				
		Control (25% methomyl)	1 % extract	2 % extract	3 % extract	4 % extract
12 hours	Leaves	3.00 ± 1.00 ^b	3.67 ± 5.77 ^{ab}	3.00 ± 0.00 ^{ab}	2.00 ± 0.00 ^{ab}	1.00 ± 0.00 ^a
	Stems	3.00 ± 1.00 ^b	0.00 ± 0.00 ^a	0.00 ± 0.00 ^a	0.00 ± 0.00 ^a	0.00 ± 0.00 ^a
24 hours	Leaves	3.67 ± 5.77 ^b	1.67 ± 0.57 ^b	2.33 ± 5.77 ^{ab}	2.67 ± 5.77 ^{ab}	1.67 ± 5.77 ^b
	Stems	5.67 ± 5.77 ^b	0.67 ± 0.57 ^a	0.67 ± 0.57 ^a	0.67 ± 5.77 ^a	0.33 ± 5.77 ^a
48 hours	Leaves	7.67 ± 5.77 ^c	3.00 ± 1.00 ^a	3.00 ± 1.00 ^{ab}	6.00 ± 1.00 ^{bc}	4.33 ± 1.15 ^{ab}
	Stems	7.67 ± 5.77 ^b	0.67 ± 0.57 ^a	0.67 ± 0.57 ^a	2.00 ± 1.00 ^a	0.33 ± 0.57 ^a
72 hours	Leaves	10.00 ± 0.00 ^c	4.33 ± 1.15 ^a	6.33 ± 5.77 ^{ab}	8.33 ± 1.53 ^{bc}	5.00 ± 1.00 ^a
	Stems	7.67 ± 1.15 ^b	1.00 ± 0.00 ^a	1.00 ± 0.00 ^a	2.00 ± 1.00 ^a	0.67 ± 0.57 ^a

Notes: Mean values of cocoa mealybug mortality for the treatment of kitolod leaf and stem extracts followed by the same letter in the same row show no significant difference at the $\alpha = 5\%$ level.

At the 72-hour observation time, the 3% concentration of kitolod leaf and stem extracts caused the highest average mealybug mortality, which was also significantly higher than the mealybug mortality caused by the 4% concentration of kitolod for leaf methanol extract but not significant for the stem methanol extracts. Overall, for each concentration treatment at a certain observation time, the death of mealybug cocoa caused by methanol extract of kitolod leaves was higher than that caused by methanol extract of kitolod stems.

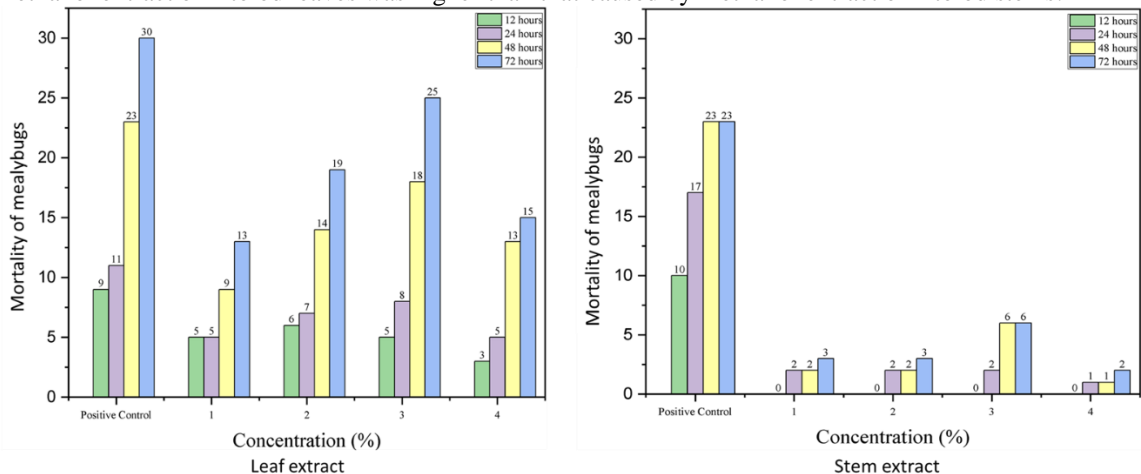


Figure 1. Number of cocoa mealybug mortality in each treatment of concentration and observation time.

From Figure 1, it can be seen that the average mortality of mealybugs due to concentration treatment and observation time shows that only the treatment of methanol extract of kitolod leaves whose extract concentration test resulted in the death of mealybugs was 50%. Thus, the LC₅₀ value was only carried out for the methanol extract of kitolod leaves (Table 4). However, the LT₅₀ value was carried out for both the leaf extract and the stem extract of kitolod (Table 5).

Table 5. LC₅₀ value of methanol extract of kitolod leaves at different observation times

Hours after treatment	Value of LC ₅₀ (%)	Std. Error	Lower Bound	Upper Bound
12 hours	0.185	0.469	-1.686	0.152
24 hours	3.065	0,45	-1.495	0.258
48 hours	3.065	0.424	-1.575	0.088
72 hours	6.016	0.491	-2.339	-0.414

Significance level for use of heterogeneity factor = 0.05
 Transformation = Log base 10

Based on the results of probit analysis, the LC₅₀ value for kitolod leaf extract is 3.068% at the observation time of 48 hours after treatment (Table 5). The LT₅₀ value for both leaf and stem extracts of kitolod was obtained from the 3% concentration treatment but with a smaller LT₅₀ value for leaf extracts which was 2.371 hours compared to stem extracts which required a longer time of 7.024 hours (Table 6).

Table 6. LT₅₀ values of methanol extracts of kitolod leaves and stems

Concentration of Methanol Extract	Source of Extract	Value of LT ₅₀ (%)	Std. Error	Lower Bound	Upper Bound
Positive control	Leaves	1.754	0.608	2.292	4.677
	Stems	1.593	0.539	1.028	3.141
1%	Leaves	6.674	0.580	0.226	2.498
	Stems	19.274	1.122	-0.406	3.992
2%	Leaves	3.122	0.563	0.968	3.174
	Stems	19.274	1.122	-0.406	3.992
3%	Leaves	2.371	0.587	1.812	4.112
	Stems	7.024	1.070	0.715	4.910
4%	Leaves	3.995	0.630	1.127	3.597
	Stems	25.108	1.448	-0.965	4.713

Logarithm base = 10.

DISCUSSION

Of all concentration treatments, the positive control caused the highest cocoa mealybug mortality (Tables 2 and 4) which was significantly higher than all kitolod methanol extract concentration treatments. However, the results of this study also showed that both methanol extracts of kitolod leaves and stems showed toxicity to cocoa mealybugs. These results prove that methanol extracts of kitolod leaves and stems contain compounds that have bioactivity as toxins against cacao mealybugs. This result is in line with the results of research by Putri et al. (13) which proved that kitolod contains alkaloid, flavonoid, tannin, quinone, polyphenolate, and saponin compounds. According to Kurniawan et al. (14), alkaloids are effective in killing mealybugs because alkaloids can attack the digestive tract of mealybugs and then inhibit their eating activity (antifeedant) as a result of damage to the digestive organs. The toxic nature of flavonoids against mealybugs is because they can function as stomach poisons (stomach poisoning) which will interfere with the digestive organs, causing the metabolic process to be inhibited. Flavonoid compounds can also cause irritation to the skin of pests. This compound enters the pest's body through the thin cuticle in the segment area and respiratory tract (spiralkulum), causing poisoning (10). Saponins are effective in killing mealybugs because saponins act as stomach poisoning which can reduce the surface tension of the mucous membrane of the pest digestive tract so that it becomes corrosive. Saponin compounds that enter the body of insect pests will change their eating

behavior by inhibiting metabolism in the digestive tract. Saponins can also inhibit the growth of larval stages by inhibiting development at the larval moulting stage (15).

The number of mealybug mortalities caused by methanolic extract of kitolod leaves was higher than that caused by methanolic extract of the stem (Tables 2, and 4) indicating that the content of active compounds in the kitolod leaf extract was higher than that contained by the kitolod stem extract, making it more toxic to mealybugs than the kitolod stem extract. Due to the low content of active compounds in the kitolod stem extract, there was no concentration of kitolod stem extract tested in this study that caused 50% mortality of mealybugs (Figure 1). This assumption is in line with the results of research by Agustrina et al. (16) which showed that bandotan dried leaf extract was more effective in reducing disease severity in chickpea plants infected with bandotan dried stem.

Flavonoids, alkaloids, and terpenes are secondary metabolites in plants that have various bioactivities. There are three main metabolic pathways of secondary metabolites in plants, namely 1) the cyclic acid pathway, 2) the mevalonic acid and methylerythritol phosphate (MEP) pathway, and 3) the malonic pathway. The main precursors of these three metabolic pathways of secondary metabolites are compounds of intermediates from carbon metabolism, namely the respiration-photosynthesis process (17) which takes place in the leaves. Thus, in addition to being a food-producing organ or energy source, namely carbohydrates, leaves also act as a place of secondary metabolite synthesis, so that the content of secondary metabolites in leaves is higher than in other organs. Atisha and mita (18) showed that the content of secondary metabolites in bandotan leaves is more diverse and for certain secondary metabolite compounds the content is higher than in stems, roots, and flowers.

The results of probit analysis showed that the LC_{50} for kitolod leaf extract was 3.065% at 48 hours after treatment (Table 5). These results indicate that the toxicity of the active compounds contained in the kitolod leaf extract is very effective in killing mealybugs and makes kitolod leaves a potential source of bioinsecticide. At a concentration of 4% kitolod leaf extract, the number of mealybug mortalities decreased. The decrease in the number of mealybug mortalities is thought to be due to the detoxification process in the mealybug digestive system, making the mealybugs more resistant to the toxins contained in the kitolod leaf extract. According to Djojosemarto (19), too high concentrations of active compounds can accelerate the occurrence of mealybug resistance to active compounds because the ability of insects to detoxify the toxins contained in these active compounds increases. On the other hand, concentrations of active compounds that are too low can cause pest tolerance to them. While Dono et al. (20) argue that the cause of the decrease in the mortality rate of mealybugs against active compounds that are stomach poisons is due to an increase in the resistance of the digestive tract wall to these active compounds. The mechanism of action of active compounds that are stomach poisons takes place by entering with the food it sucks through the test media that has been given extracts containing active compounds. Toxins that enter the stomach are then absorbed by the intestinal wall and then attack their targets according to the toxic properties of the active compounds (21).

Both leaf and stem extracts of kitolod showed LT_{50} at 3% extract concentration (Table 6), but with a faster death time in leaf extracts, which is about 2.371 hours compared to death caused by stem extracts which takes up to 7.024 hours. The results are consistent with the assumption that the content of active compounds that are toxic in the kitolod leaf extract is higher than in the kitolod stem extract. These results also provide assurance that kitolod leaves have great potential as a source of bioinsecticides that can replace the use of chemical insecticides that are not environmentally friendly for plant pest control.

CONCLUSION

Based on the results and discussion above, it is known that both leaf and stem methanol extracts show their toxicity, causing death to mealybugs of cocoa plants. Based on the LC_{50} and LT_{50} values obtained, it is known that:

1. LC_{50} of methanol extract of kitolod leaves was 3.065% while methanol extract of kitolod stems showed no LC_{50} .
2. The LT_{50} of methanol extract of kitolod leaves and stems was obtained at 3% extract concentration with smaller LT_{50} value of leaf extract compared to kitolod stem extract.

Thus, it can be concluded that the toxicity of methanol extract of kitolod leaves on cocoa plant mealybugs is higher than that of methanol extract of kitolod stems.

REFERENCES

- [1] J. Milz, B. Regine, W. Nurheni, A. Afwan, dan T. Heinrich. *Pengelolaan Kebun Kakao Berkelanjutan*. Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) Gmb., Jakarta, 2016).
- [2] A. F. K. Wardani. “Karakterisasi Morfologi Tanaman Kakao (*Theobroma cacao* L.) Hibrida F1 Lindak di Wisata Edukasi Kampung Coklat Blitar Sebagai Sumber Belajar Biologi”, Skripsi, IAIN, Tulungagung, 2019.
- [3] Badan Pusat Statistik, “Produksi Tanaman Perkebunan” (BPS, Jakarta, 2020)
- [4] M. Mastura dan N. Nuriana, *Jurnal Pendidikan Kimia dan Ilmu Kimia*. 1 (2), 29-36. (2018).
- [5] F. Arfianto, *Jurnal Daun*. 5(1): 17- 26. (2018).
- [6] A. Fitri. “Pengaruh Ekstrak Biji Buah Srikaya (*Annona Squamosa*) Terhadap Hama Kutu Putih *Pseudococcus viburni*”, Skripsi, UIN Raden Intan, Bandar Lampung 2019.
- [7] R. Ramadhona, D. Djamilah., M. Mukhtasar. *Jurnal Ilmu-Ilmu Pertanian Indonesia*. 20(1):1-6 (2018).
- [8] N. Nukmal, dan R. Andriyani. “Daya Insektisida Ekstrak Polar Serbuk Daun Gamal Kultivar Pringsewu Terhadap Kutu Putih (Hemiptera: Pseudococcidae) pada Kakao” in *Pertanian dan Tanaman Herbal Berkelanjutan di Indonesia*. Prosiding Seminar Nasional. (UMJ, Jakarta, 2017), pp. 127-137.
- [9] E. Malik, dan M. Dewi. *Jurnal Farmasetis*. 3(2): 37-41 (2014).
- [10] N. Nukmal, G. D. Pratami, E. Rosa, A. Sari, and M. Kanedi. *Journal of Agriculture and Veterinary Science*. 12(3): 4-8 (2019).
- [11] R. Rusman, R. F. I. Rahmayani, dan M. Mukhlis. *Buku ajar kimia larutan* (Syiah Kuala University Press. 2020).
- [12] N. Novriadi. “Kuantifikasi dan Karakterisasi Senyawa Flavonoid Ekstrak Polar Daun Gamal Kultivar Bandar Lampung untuk Mengetahui Daya Toksisitasnya Terhadap Hama Kutu Putih”. Thesis, Magister Biologi, Universitas Lampung, Bandar Lampung 2019.
- [13] D. D. Putri, S. Hazar, dan S. P. Fitrianiingsih. “Antibacterial Activity Test of Kitolod (*Isotoma longiflora* (L.) C. Presl.)”, *Prosiding Farmasi*, (Bandung, 2016), pp. 529-535.
- [14] A. Kurniawan, M. Muhfahroyin, dan A. Sutanto. *Biolova*. 2(1): 54-63 (2021).
- [15] R. Eka, M. Moerfiah, dan Triastinurmiatiningsih. *Jurnal Ilmiah Ilmu Dasar dan Lingkungan Hidup*. 18(2): 55-62 (2018).
- [16] R. Agustrina, N. Fadhillah, D. E. Tamara, L. Chrisnawati, dan Yulianty, *Pemanfaatan Bandotan.*, *Jurnal Biodjati*, in press.
- [17] B. M. Twaij and M. N. Hasan. . *Plant Biol*. 13(1), 4-14 (2022).
- [18] S. A. Atisha dan S. R. Mita. *Farmaka*. 16 (3): 116 -121 (2018).
- [19] Djojsumarto, P. 2008. *Pestisida dan Aplikasinya*. Jakarta Selatan: PT Agromedia Pustaka
- [20] D. Dono, S. Ismayana, I. Idar, D. Priyono, dan I. Muslikha I. *Jurnal Entomologi Indonesia*. 7(1):9-27 (2010).
- [21] R. Tarumingkeng, R. *Insektisida; Sifat, Mekanisme, Kerja dan Dampak Penggunaannya* (Ukrida Press, Jakarta, 1992)

Analysis of Coverage and Distribution of *Melastoma malabathricum* Invasion in Way Kambas National Park

Jani Master^{*(1)}, Santoso⁽²⁾, Ahmad Fanani⁽²⁾, Sumianto⁽²⁾, Ichan Prastika⁽²⁾,
Muhammad Yunus⁽²⁾

⁽¹⁾Department of Biology, Faculty of Mathematics and Natural Sciences, Universitas Lampung, Indonesia

⁽²⁾Yayasan Penyelamatan dan Konservasi Harimau Sumatera, Indonesia

*Corresponding author: j.janter@gmail.com

Abstract. Way Kambas National Park (WKNP) is a important habitat for various protected species, including the Sumatran tiger, Sumatran rhinoceros, and Sumatran elephant. This conservation area faces the threat of invasion by *Melastoma malabathricum*, which has the ability to transform wetlands into dry land. This study aims to map and quantify changes in the coverage of *M. malabathricum* in WKNP from 1993 to 2019. The method used is the analysis of land cover with Landsat imagery covering the Way Kambas National Park area from 1993 to 2019 using QGIS 2.6.0 software with the Molusce plugin. The results showed changes in *M. malabathricum* coverage from 1993 to 2019, with the largest increase occurring between 1993 and 1998, reaching 15.961,23 hectares. The majority of the invasion occurred in open areas.

Keywords: invasive species, conservation area, land coverage

INTRODUCTION

Sumatra is one of the islands in Indonesia that has a high biodiversity. Several large mammals are protected here, such as the Sumatran rhinoceros, Sumatran elephant, and Sumatran tiger [1]. These animals are distributed across various locations on the island, one of which is the Way Kambas National Park (WKNP). Therefore, WKNP has been designated as a habitat for several priority conservation species, including the sumatran tiger (*Panthera tigris sumatrae*), sumatran rhinoceros (*Dicerorhinus sumatrensis*), Sumatran elephant (*Elephas maximus sumatrensis*), tapir (*Tapirus indicus*), sun bear (*Helarctos malayanus*), crested fireback (*Lophura ignita*), great argus (*Argusianus argus*), white-winged wood duck (*Asarcornis scutulata*), various species of eagles, various species of hornbills, various species of kingfishers, and various species of pitcher plants through Minister of Forestry Regulation Number P.57/Menhut-II/2008.

The majority of WKNP is comprised of secondary forests, shrublands, grasslands, and wetlands [2]. It represents the largest lowland forest area in Sumatra [3] [4]. The abundance of natural resources in WKNP is not without its challenges in management. Some of the issues faced in WKNP include encroachment, hunting, and forest fires. Forest fires occur almost every year in WKNP, especially during the dry season [4]. These fires have altered the vegetation conditions in WKNP. In its history, certain areas in WKNP, such as Rawa Kalibiru, used to be freshwater swamp forests dominated by trees adapted to waterlogging. However, due to recurring fires almost every year, the swamp transformed into grasslands [5]. Over time, these grasslands have become crucial areas for the survival of wildlife in Way Kambas National Park, including sambar deer, sumatran elephants, wild boars, and various species of water birds.

Currently, the grassland condition of Rawa Kalibiru is gradually shrinking, replaced by the dominance of *Melastoma malabathricum* L. plants [6]. Plant invasion has become a global concern due to the negative impacts caused by the uncontrolled growth of these plants in specific locations, resulting in environmental damage and a decline in biodiversity [7]. The Ministry of Environment, in collaboration with SEAMEO BIOTROP, has reported at least 1,936 invasive plant species, and *M. malabathricum* is among the 75 most threatening invasive species [8] [9]. The invasion of *M. malabathricum* in WKNP is not only limited to Rawa Kalibiru but also occurs in several other locations. Therefore, mapping the distribution of these plants and monitoring changes in their coverage periodically is necessary. This information can serve as input for the Way Kambas National Park management scheme in determining area management strategies because there is currently no information regarding the distribution of *M. malabathricum* invasion in TNWK.

MATERIALS AND METHODS

Distribution Analysis of Melastoma malabatricum

The distribution of *M. malabatricum* can be determined by conducting spatial analysis on Landsat imagery covering the Way Kambas National Park area. The steps involved are as follows [10] :

1. Geometric correction of Landsat imagery
This step involves projecting the map data into a specific map projection system to reduce geometric distortions. The result of this correction is the Root Mean Square (RMS) value, with a maximum value of 0.5. The correction becomes more accurate by increasing the number of Ground Control Points (GCP) and discarding RMS values greater than 0.5.
2. Image cropping according to the research boundary
This step is performed to focus on the area that is the subject of land cover analysis.
3. Supervised classification
Supervised classification involves grouping land cover classes based on their characteristic signatures obtained through the creation of training areas. The supervised classification is performed using Quantum GIS 2.6.0 software with the Semi-Automatic Classification Plugin (Molusce plugin). The land cover classes classified include shrubs, *Melastoma malabatricum* communities, dryland forests, swamp forests, mangroves, water bodies, and marshes.
4. Validation
After the supervised classification is completed, the next step is to perform Accuracy Assessment (validation) by comparing the field data, in the form of coordinate points, with the classified land cover. The classification validation is expected to have a value of over 85%, indicating the validity of the land/vegetation cover classification.

Land Cover Change Analysis

Land cover change involves comparing the land cover conditions of a specific area between two periods of time. In this research, the focus is on studying the changes in land cover to examine the invasion of the *M. malabarrichum* species from 1993 to 2019 with a 5-year interval. The analysis of land cover change is conducted using QGIS 2.6.0 software with the Molusce plugin. To observe changes in the distribution of *M. malabathricum*, it is done by comparing the *M. malabathricum* coverage conditions for each year analyzed, which are the years 1993, 1998, 2003, 2008, 2013, 2019 (Table 1).

Table 1 The source of data for **land** cover change and *M. malabarrichum* coverage in Way Kambas National Park

Year	Data source	Data recording time	
		Path/Row 123/63	Path/Row 123/64
1993	Landsat 5 Path/Row 123/63 and 123/64	March 18, 1993	August 9, 1993
1998	Landsat 5 Path/Row 123/63 and 123/64	March 16, 1998	December 13, 1998
2003	Landsat 5 Path/Row 123/63 and 123/64	November 25, 2003	March 5, 2003
2008	Landsat 5 Path/Row 123/63 and 123/64	March 27, 2008	December 24, 2008
2013	Landsat 8 OLI/TIRS Path/Row 123/63 and 123/64	April 29, 2013	April 29, 2013
2019	Landsat 8 OLI/TIRS Path/Row 123/63 and 123/64	July 21, 2019	July 21, 2019

RESULTS AND DISCUSSION

The results of image interpretation for the years 1993, 1998, 2003, 2008, 2013, and 2019 show changes in the extent of each land cover class. Data on land cover are presented in Table 2. Forests dominate the land cover in each year. Forest cover experienced a decrease in 1998, 2003, and 2008, which is consistent with the land cover study conducted by Maulana and Darmawan [2], which reported a decrease in forest cover in 1996, 2002, and 2010. Forest cover in TNWK decreased by around 2.6% per year [11]. There are differences in the land cover extent figures compared to previous studies conducted by Maulana and Darmawan [2] and

Andyono *et al* [11] due to differences in methods and sources of Landsat imagery, which can yield different results [12] [13].

According to [14], factors contributing to land cover changes include population growth, livelihoods, accessibility, and supporting facilities of life, as well as government policies. Another factor that can cause such changes is forest fires that occurred in 1997-1998 [4]. The result of these fire incidents is often replaced by shrubland cover. When the original and natural climax vegetation, which is the forest, is damaged through logging, shifting cultivation, or fires, it is often replaced by shrubland.

Table 2. Changes in land cover area in Way Kambas National Park from 1993 to 2019

Types of Land Cover	Area (ha)					
	1993	1998	2003	2008	2013	2019
Shrubs	35,463.05	44,161.31	22,460.00	28,400.49	36,242.82	27,231.88
Forest	64,105.73	64,924.48	50,594.81	44,385.22	63,787.3	53,878.10
Swamp forest	5,525.26	4,379.06	21,921.34	17,307.02	15,989.86	13,659.29
Water bodies	1,230.92	1,333.82	1,750.95	939.3	1,096.86	2,633.98
Swamp	7,171.14	1,981.24	7,439.16	6,107.05	4,170.01	5,030.45
<i>Melastoma malabatarichum</i> invasion	2,847.08	9,292.47	15,961.23	11,490.33	12,514.62	12,077.07
No Data	12,626.71	2,897.50	8,842.49	20,340.47	4,680.22	14,459.13

In this study, we included *M. malabatarichum* as a separate land cover class to analyze its distribution and invasion changes from 1993 to 2019. Difficulties arose when distinguishing between shrubs and *M. malabatarichum* cover. *M. malabatarichum* is an invasive species with a wide distribution in Way Kambas National Park. The highest increase in *M. malabatarichum* invasion occurred during the period of 1993-1998, reaching 69%. However, there was also a decrease in *M. malabatarichum* invasion during the years 2003-2008, with a decrease of 39% (Figure 1).

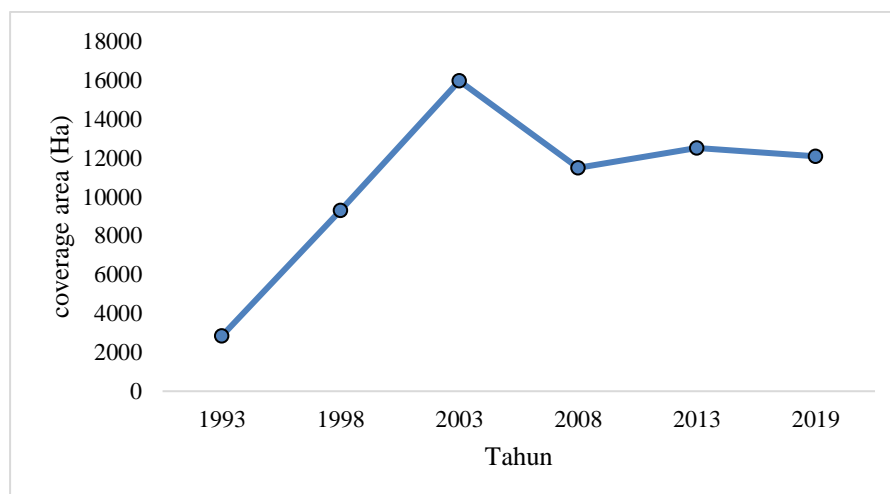


Figure 1. Changes in the area of *M. malabatarichum* from 1993 to 2019

Changes in *M. malabatarichum* may also be influenced by seasonal conditions during satellite image acquisition. For example, the Landsat 5 Path/Row 123/63 image taken in November, which is the end of the dry season in WKNP, shows an area that was classified as water bodies on the 1998 and 2008 maps but filled with *M. malabatarichum* on the 2003 map. The dry season in WKNP can lead to drought conditions in the marshes within the national park. The cause of *M. malabatarichum* invasion is still being studied, whether it is due to forest fires that often occur in WKNP or other contributing factors. However, generally, *M. malabatarichum* invasion occurs in open forest conditions (Figure 2).

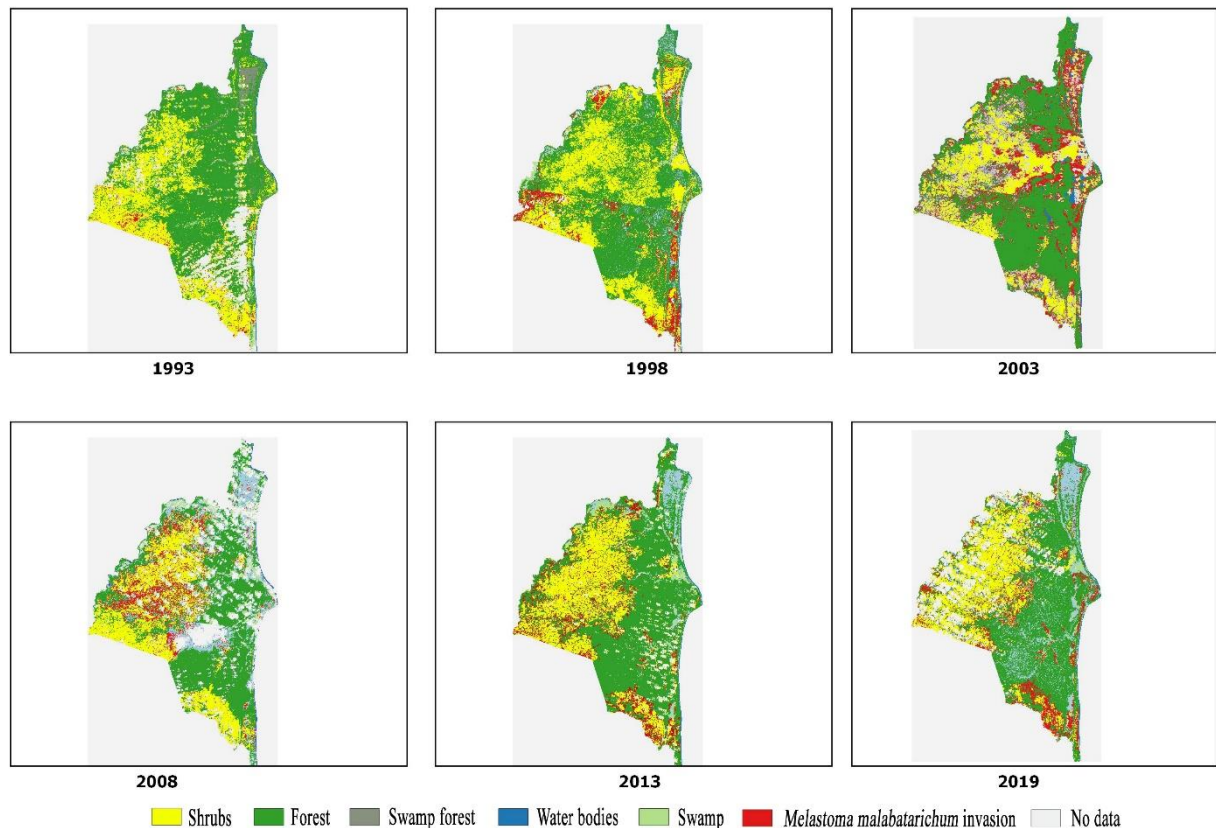


Figure 2. Changes in land cover in Way Kambas National Park from 1993 to 2019

CONCLUSION

M. malabatarichum is distributed in open locations within WKNP. There have been an increase in the cover of *M. malabatarichum* from 1993 to 2019, with the largest increase occurring during the period of 1993-1998.

ACKNOWLEDGEMENT

A gratitude is expressed to Way Kambas National Park, Yayasan Penyelamatan dan Konservasi Harimau Sumatra (PKHS) of Way Kambas National Park, Lampung, Indonesia and Tropical Forest Conservation Action for Sumatra (TFCA-Sumatra) for facilitating this research so that this research can proceed as expected.

REFERENCES

- [1] M. I. Lubis, J. S. Lee, U. M. Rahmat, E. R. Tarmizi, E. Ramadiyanta, D. Melvern, S. Suryometaram, A. Trihangga, M. Isa, D. Yansyah, R. Abdullah, Ardiantiono, W. Marthy, K. R. Jones, N. Andayani, M. Linkie, . Planning for Megafauna Recovery in the Tropical Rainforests of Sumatra. *Frontiers in Ecology and Evolution* 11: 1-14 (2023).
- [2] J. Supriatna, C. Margules, *The National Parks of Indonesia* (Yayasan Pustaka Obor Indonesia, Jakarta, 2022)
- [3] Balai Taman Nasional Way Kambas, *Laporan Tahunan Taman Nasional Way Kambas Lampung* (Balai Taman Nasional Way Kambas, Lampung, 2013).
- [4] G. D. Winarno, S. P. Harianto, B. S. Dewi, Landscape characteristics analysis of Rawa Bunder Resort at Way Kambas National Park. *Journal of Natural Resources & Environment Management* 13(2) : 199-210 (2023)

- [5] Directorate General of Forest Protection and Nature Conservation. Way Kambas National Park management Plant 1994-2019 Volume II: Data Projections and Analysis (Word Bank, Bogor, 1995).
- [6] J. Master, I. Qoyim, D. Setiadi, N. Santoso, Autecology of *Melastoma malabathricum*, an invasive species in the Way Kambas National Park, Indonesia. *Biodiversitas* 21: 2303-2310 (2020).
- [7] J. Master, S Tjitrosoedirdjo, I. Qoyim, S.S. Tjitrosoedirdjo, Ecological impact of *Merremia peltata* (L.) Merrill invasion on plant diversity at Bukit Barisan Selatan National Park. *Biotropia* 20(1): 29-37 (2013).
- [8] S. Tjitrosoedirdjo, S.S. Tjitrosoedirdji, T. Setyawati, Tumbuhan Invasif dan Pendekatan Pengelolaannya. (Seameo Biotrop, Bogor, 2016).
- [9] S.S. Tjitrosoedirdji, I. Mawardi, S. Tjitrosoedirdji, 75 Important Invasive Plant Species in Indonesia (Seameo Biotrop, Bogor, 2016)
- [10] L. Congedo, Semi-Automatic Classification Plugin for QGIS. Rome: Sapienza University. Available at: <http://www.planning4adaptation.eu/> (2013).
- [11] G. Andyono, D. Marsono, R. Sadono, M.A. Imron, Analisis Penginderaan Jauh Multi-Temporal Terhadap Perubahan Penutupan Lahan Di Daerah Penyangga Dan Dalam Kawasan Taman Nasional Way Kambas, Lampung. *J. Penelit. Kehutan. Sumatera* 2(1):1-11 (2018).
- [12] P. Gong, P.J. Howarth, An assessment of some factors influencing multispectral land-cover classification. *Photogrammetric Engineering and Remote Sensing* 56(5), 597-603 (1990).
- [13] N.A. Mahmon, N. Ya'acob, A.L. Yusof, Differences of Image Classification Techniques for Land Use and Land Cover Classification. (In IEEE Transactions on Geoscience and Remote Sensing, 2015), pp. 6-8.
- [14] P. Amalina, L.B. Prasetyo, S.B.Rushayati, Forest Fire Vulnerability Mapping in Way Kambas National Park, *Procedia Environmental Sciences* 33, 239-252 (2016).

Morphological Variation of Climbing Perch (*Anabas testudineus* Bloch 1972) in South Kalimantan Based on Morphometric Characters

Rinta Dwi Takarini^{*(1)}, Rani Sasmita⁽²⁾, Badruzaufari⁽³⁾

⁽¹⁾Department of Biology, Faculty of Mathematics and Natural Sciences, Universitas Lambung Mangkurat

*Corresponding author: *1911013220005@mhs.ulm.ac.id

Abstract. Climbing perch (*Anabas testudineus* Bloch 1972) is a freshwater fish that is classified as having a high tolerance for environments with acidic water conditions. The rivers and swamps in South Kalimantan are known to have low pH levels. Therefore, information is needed regarding morphological characters for management of Climbing perch resources. This information can be used to identify population units in South Kalimantan. The research method used was Principal Component Analysis (PCA) based on morphometric characters in Climbing perch. Sampling was carried out in South Kalimantan, Danau Panggang, Margasari, Sungai Tabuk, Batakan and Kurau. Of the 33 morphometric characters observed, there were 17 morphometric characters that contributed to influence the results of the Principal Component Analysis. We suggested this research will continue with a molecular approach to determine whether or not there is an influence of morphological variations on genetic changes in fish.

Keywords: papuyu, clustering, phenetic, morphometric

INTRODUCTION

South Kalimantan has abundant freshwater fishery potential which is scattered in every district and city, one of which is climbing perch (*Anabas testudineus* Bloch 1972). Climbing perch has the potential to be an alternative food source to prevent malnutrition because every 100 grams of fresh climbing perch meat contains 4.6 g carbohydrates, 14.3 g protein, 4.9 g fat, 329 mg calcium, 1.1 mg Zn (Nurhamidah & Dara, 2022). According to Mohanty et al. (2016), climbing perch contains 2 groups of omega 3 which are important for brain nutrition, namely DHA (Docosahexaenoic Acid) and EPA (Eicosapentaenoic Acid) whose concentration is equivalent to Indian oil sardine fish (*Sardinella longiceps*). The taste of meat is tasty and slightly fatty, making climbing perch very popular with the public (Mustika et al., 2020), even though the price is high, reaching Rp. 70,000.00-120,000.00 per/kg (Slamat et al., 2017). The economic value of climbing perch, which is relatively high, causes demand in the market to increase and encourages an increase in fishing for papuyu in nature by fishermen (Slamat et al., 2012). According to Salamat et al. (2017), the need for climbing perch in South Kalimantan reaches >800 kg/day, which is almost entirely taken from the availability in nature, while only 10% is taken from cultivation. Based on data on the volume of PUD of climbing perch in South Kalimantan in 2021 it was 12,072 tons, higher than in 2020 which amounted to 11,095 tons, while aquaculture production was only 132 tons in 2021 (KKP Statistics, 2023).

The high rate of catching Climbing perch is not accompanied by aquaculture production. This is feared to cause the presence of climbing perch in nature to decrease so that there will be scarcity and damage to swamp water ecosystems as a result of fishing processes that are not environmentally friendly such as the use of strum, chemicals, and non-selective fishing gear (Slamat et al., 2017). According to Mustika et al. (2020), the conversion of swamp land to plantations is the cause of the decline in climbing perch populations in nature. Even though this type of climbing perch has a dominant habitat in swamp waters such as monotonous, tidal, and rainfed swamps (Slamat et al., 2012).

Threatened the existence of climbing perch can cause habitat differences and make exists difference morphology between papuyu with the others in each area. For it is necessary something an effort properly conservation for preserving climbing perch. The stage at first is through approach morphology. Identification with a morphological approach is an important part of efforts to manage inland public waters fisheries resources that are threatened with extinction. Research on the study of morphological characteristics of climbing perch in South Kalimantan is still limited while catching climbing perch will continue. Therefore, research is needed to identify the morphometric characters of climbing perch (*Anabas testudineus*) in South Kalimantan. This research is expected to be a valuable input as a basis for the management of climbing perch production in conservation as well as data in determining climbing perch conservation policies in the future.

MATERIALS AND METHODS

This research was conducted in March-May 2023 at the Genetics and Molecular Laboratory of the Faculty of Mathematics and Natural Sciences, University of Lambung Mangkurat Banjarbaru. The series of studies conducted at the Laboratory included observing and measuring fish samples taken at five locations, namely Danau Panggang (Hulu Sungai Utara District), Candi Laras Utara (Tapin District), Lok Baintan (Sungai Tabuk District), Kurau (Tanah Laut Regency), and Batakan (Tanah Laut Regency).

The materials used for fish observation were water, ice, and 96% alcohol as a fixative solution. The tools used, namely calipers, rulers, cool boxes, stationery, latex, cellphone cameras, and cutting mats.

The method used in this study is the observation method, which will produce primary data. The primary data was obtained from measurements of fish samples in the laboratory. Then, the primary data was analyzed using PAST software version 4.03 to get the results of clustering climbing perch from five locations in South Kalimantan. Then, to find out the main differentiating factors from morphological characters, Principal Component Analysis is used. Furthermore, identification was carried out using a molecular approach to ensure that the fish sample used was *Anabas testudineus*.

Calculation Method of Morphometric Characters of Climbing Perch (*Anabas testudineus*)

Samples obtained from the field were brought to the Genetics and Molecular Laboratory for morphometric measurements. Before fish are measured, documentation and labeling are carried out. The morphometric character measurements are 33 characters and refer to the morphometric measurement data following Akbar's instructions (2018). Previously measured fish samples were frozen and stored in the freezer. The measured body parts can be seen in **Figure 1**.

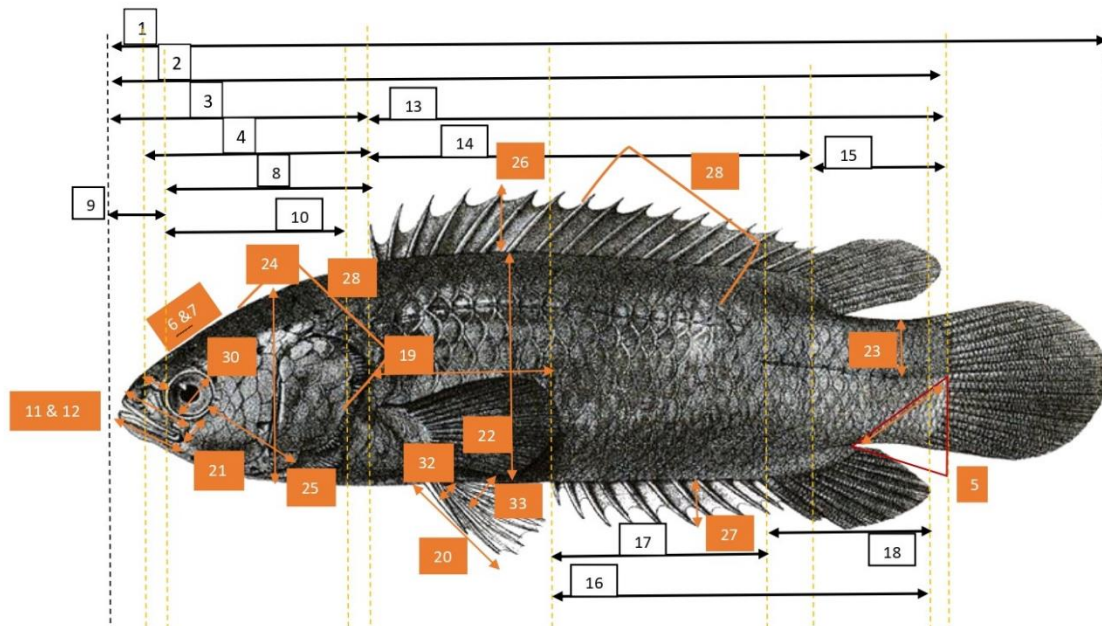


Figure 1. Measurements Morphometrics of Climbing perch (*Anabas testudineus*)

Description :

1). Total length (PT), 2). Standard length (PB), 3). Head length (PK), 4). Long ahead dorsal fin (PSD), 5). stem length tail (PBE), 6). Nose length (PH), 7). space length between eyes (PRAM), 8). Head length behind _ eyes (PKBM), 9). Head length in front eyes (PKDM), 10). Intermediate length eye with preoperculum (PAMP), 11). Jaw length above (PRA), 12). Jaw length bottom (PRB), 13). Dorsal fin length (PDSD), 14). Base length fingers hard dorsal fin (PJKD), 15). Base length fingers weak dorsal fin (PJLD), 16). Base length anal fin (PDSA), 17). Length of fingers hard anal fin (PJKA), 18). Length of fingers weak anal fin (PJLA), 19). fin length pectorals (PSP), 20). Ventral fin length (PSV), 21). High below eyes (TBM), 22). Height (TB), 23). Stem height tail (TBE), 24). Head height (TK), 25). Cheek height (TP), 26). Dorsal fin height (TSD), 27). Anal fin height (TSA), 28). Body width (LB), 29). Head width (LK), 30). Eye width (LM), 31). Aperture width

mouth (LBM), 32). Base length fingers hard ventral fin (PJKV), 33). Base length fingers weak ventral fins (PJLV).

RESULTS AND DISCUSSION

Morphological Observation Results of Climbing Perch (Anabas testudineus)

The research results on the climbing perch's morphological characteristics, having a blunt and scaly head, a short snout, and a terminal mouth shape. Climbing perch lips thin and jagged. It has a flat, upright, and rounded body. It has a dorsal fin which is located on the anterior of the body. It has a separate dorsal fin from the caudal fin. The climbing perch has a type of gray to greenish color, and the tips of the scales and fins are brightly colored. Has a serrated section on the operculum and preopercle. The body model of the climbing perch bends inwards. These characteristics are to previous research by Akbar (2018). Images of climbing perch observed in this study are presented in Figure 2.

Clustering Results of Climbing Perch (Anabas testudineus)

The clustering results in this study were taken from data measuring morphometric characters and are generally depicted with a tree diagram or commonly known as a dendrogram. The clustering results show that there are groupings that are uneven and tend to be parallel to the dendrogram line, except for the samples from the Tabuk and Kurau River locations. The number of clusters formed can be determined from the dendrogram that occurs and depends on the subjectivity of the researcher (Luthfi & Wijayanto, 2021). The dendrogram is presented in **Figure 3**.

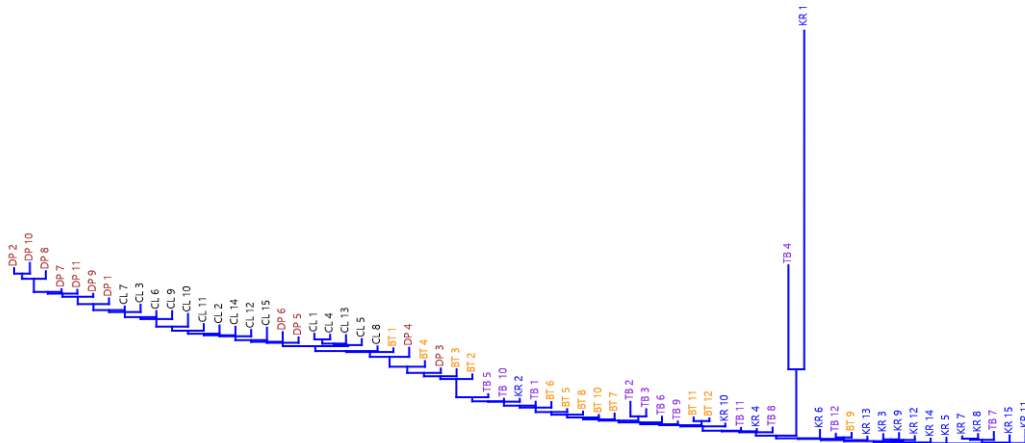


Figure 3. Results of Climbing Perch Clustering from Five Locations in South Kalimantan.

Principal Component Analysis Climbing Perch (PCA)

Table 3. Morphometric Characters and *Eigen Value*

Characters	<i>Eigen Values</i>
Long base of ventral fin weak rays (PJLV)	0.52329
Long base of ventral fin spines (PJKV)	0.39411
Mouth opening width (LBM)	0.44338
Long base of dorsal fin weak rays (PJLD)	0.3224
Eye width (LM)	0.2567
Maxilla length (PRA)	0.25301

Head Height (TK)	0.10379
Cheek height (TP)	-0.26199
Ventral fin length (PSV)	0.2824
Total length (PT)	0.3029
Head length behind eyes (PKBM)	0.28831
Height (TB)	0.2925
Standard length (PB)	0.303
head length (PK)	0.29799
The length of the base of the dorsal fin spine (PJKD)	0.29587
The length of the base of the anal fin (PDSA)	0.29213
Head width (LK)	0.30098

Based on the data in Table 3, the Eigen Value (>0.2) obtained from the Principal Component Analysis, the morphometric character that contributes to the highest influence is the base length of the weak ventral fin radius (PJLV), with an Eigen Value of 0.52329, followed by with the width of the mouth opening (LBM) with an Eigen Value of 0.44338, followed by the basic length of the ventral fin rays (PJKV) with an Eigen Value of 0.39411, the basic length of the dorsal fin soft rays (PJLD) with an Eigen Value of 0.3224, the length standard (PB) with an Eigen Value of 0.303, followed by Head Width (LK) with an Eigen Value of 0.30098, Total Length (PT) with an Eigen Value of 0.3029. Head length (PK) with an Eigen Value of 0.29799, Base length of dorsal fin rays (PJKD) with an Eigen Value of 0.29587, Body height (TB) with an Eigen Value of 0.2925, Anal fin base length (PDSA) with an Eigen Value of 0.29213, a Head length behind the eye (PKBM) with an Eigen Value of 0.28831, Ventral fin length (PSV) with an Eigen Value of 0.2824, Cheek height (TP) with an Eigen Value of -0.26199, Eye width (LM) with an Eigen Value of 0.2567, Maxillary Length (PRA) with an Eigen Value of 0.25301 and Head Height (TK) with an Eigen Value of 0.10379.

Climbing perch is a type of potamodromous fish, which means it can move from place to place if the freshwater ecosystem is in bad environmental conditions (Khatun et al., 2019) The habitat of the climbing perch is widespread in various types of swamps, such as monotonous swamps, rain-fed, and tides. The diverse habitats and wide distribution are thought to have caused differences in morphological characters, especially the morphometric characters in the climbing perch. This is a form of adaptation to climbing perch (Faradiana et al., 2018)

A climbing perch is a fish that lives in the waters at the bottom of the bottom, so it requires ventral fins that can keep the body in a stable condition (Prado et al., 2016). By the results of the analysis, the basic length of the ventral fin soft rays (PJLV) and the basic length of the ventral fin hard rays (PJKV) are characters that play a major role in influencing the results of the principal component analysis. The habit of the climbing perch tends to maintain its position in silence, so the dorsal fin on the climbing perch must be a barrier that functions to protect the position of the climbing perch to remain stable. This is consistent with the results of the analysis, that the base length of the dorsal fin rays (PJKD) and the base length of the dorsal fin weak rays (PJLD) have a major influence on the results of the principal component analysis. The next morphometric character that influences the results of the principal component analysis is anal fin base length (PDSA). This right is related to the stability of the fish's position, but can also affect the ability to move in the climbing perch.

The results of the principal component analysis show that the morphometric character of the width of the mouth opening in the climbing perch has a large influence contribution, this can be related to the fast movement of the climbing perch to move its mouth to reach and close its prey. The next characteristic that contributes to the influence of the results of the principal component analysis is the morphometric character of head width. The length and width of the head can be related to the size of the food consumed by the Papuan fish (Prado et al., 2016). All morphometric characters analyzed are of course related to the total length (PT) and standard length (PB) characters, this is because the total length and standard length can be used as a reference in viewing the results of the analysis on each morphometric character of climbing perch.

The total length and standard length characters can be used as a reference for growth patterns in fish. The Scatter Plot results from the Principal Component Analysis results are presented in **Figure 4**.

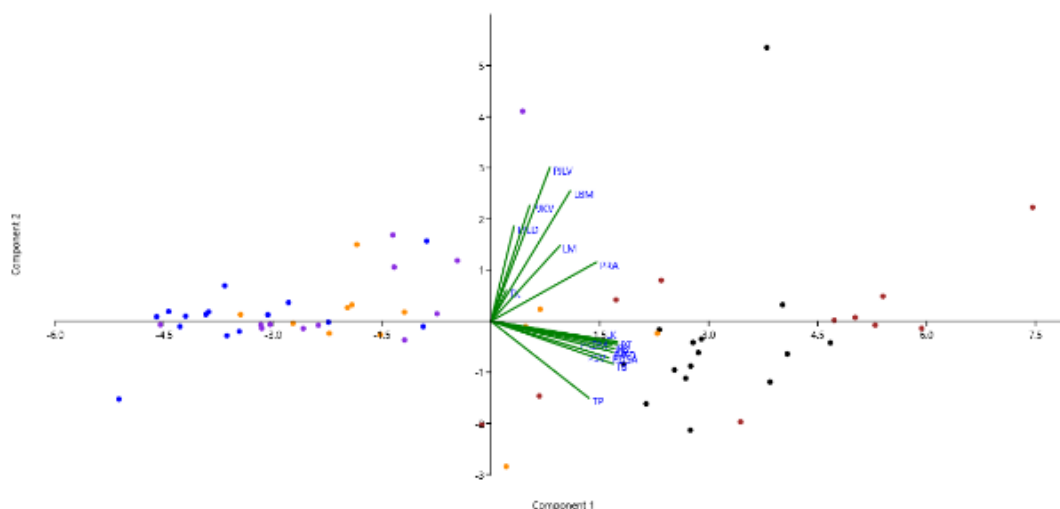


Figure 4. Scatter Plot Analysis Results Main Components

Description:

1. Blue: Kurau
2. Purple: Sungai Tabuk
3. Orange: Batakan
4. Black: Candi Laras
5. Brown: Danau Panggang

CONCLUSION

Based on the results of the Main Component Analysis, then:

1. In quadrant II, fish grouping consists of fish samples from Candi Laras, Batakan, Sungai Tabuk, and Danau Panggang. The morphometric characters that contribute to the influence, namely PSLV (base length of the weak ventral fin rays), PJKV (base length of the ventral fin rays), LBM (Width of the mouth opening), PJLD (base length of the dorsal fin weak rays), LM (eye width), and PRA (maxillary length), and TK (head height).
2. In quadrant III, fish grouping consists of fish samples from Batakan, Danau Panggang, and Candi Laras. Morphometric characters that contribute to influence, namely TP (cheek height), PSV (ventral fin length), PT (total length), PKBM (head length behind the eyes), TB (body height), PB (standard length), PK (Length of head), PJKD (Length of base of dorsal fin rays), PDSA (Length of base of anal fin), LK (Width of head)
3. In quadrants I and IV it does not show any morphometric characters, this can be interpreted that the morphometric characters in quadrants I and IV only contribute a little, so it means that the values are in contrast to the values in the characters in quadrants I and IV.
4. The results of the Clustering analysis are not well grouped, so the information obtained cannot be linked to the results of the Principal Component Analysis.

ACKNOWLEDGEMENT

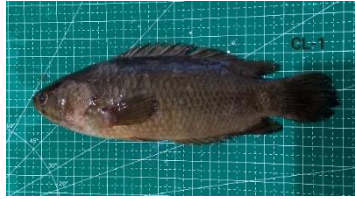
The authors would like to thank you for funding this research supported by the DIPA Universitas Lambung Mangkurat for Program Dosen Wajib Meneliti (PDWM) 2023 Nomor: SP-DIPA SP-DIPA-023.17.2.677518/2023. The authors would like to thank many parties who have assisted in this research, including supervisors, family, relatives, and friends.

REFERENCES

- [1] Akbar, J. *Ikan Papuyu Teknologi Manajemen dan Budi Daya*. (Universitas Lambung Mangkurat Press, Banjarmasin, Kalimantan Selatan, 2018).
- [2] Faradiana, R., Budiharjo, A., & Sugiyarto, S. Keanekaragaman dan Pengelompokan Jenis Ikan di Waduk Mulur Sukoharjo, Jawa Tengah, Indonesia. *Depik*, 7(2), 151-163. (2018).
- [3] Khatun, D., Hossain, M. Y., Rahman, M. A., Islam, M. A., Rahman, O., Kalam Azad, M. A., ... & Hossain, M. A. Life-History Traits of the Climbing perch *Anabas testudineus* (Bloch, 1792) in a Wetland Ecosystem. *Jordan Journal of Biological Sciences*, 12(2). (2019).
- [4] Mohanty, B. P., S. Ganguly, A. Mahanty, V. Sankar, T. R. Anandan, K. Chakraborty, ... & N. Sridhar. DHA and EPA Content and Fatty Acid Profile of 39 Food Fishes From India. *BioMed research is international*. 2016; 1-14. (2016).
- [5] Mustika, R., L. A. Sofia & E. Agusliani.. Kinerja dan Prospek Usaha Pembenuhan Ikan Papuyu (*Anabas Testudineus*) di Kalimantan Selatan. (2020).
- [6] Narita, V., A. L. Arum, M. S. Isaeni, N. Y. Fawzya. Analisis Bioinformatika Berbasis WEB untuk Eksplorasi Enzim Kitosanase Berdasarkan Kemiripan Sekuens. *Jurnal Al-Azhar Indonesia Seri Sains dan Teknologi*. 1; 197-203. (2012).
- [7] Nurhamidah, N., & W. Dara. Analysis of the Nutritional Value of Puyu Fish (*Anabas testudineus*) and Acceptability of the Processing. *Sainstek: Jurnal Sains dan Teknologi*. 14; 46-50. (2022).
- [8] Prado, A. V., Goulart, E., & Pagotto, J. Ecomorphology and use of food resources: inter-and intraspecific relationships of fish fauna associated with macrophyte stands. *Neotropical Ichthyology*, 14. (2016).
- [9] Slamant, R. R., R. K. Rina, & P. Ansyari. Pemuliaan Ikan Papuyu (*Anabas testudineus*) dengan Teknik Hybridisasi Filogenetik Meristik Dari Tiga Tipe Ekosistem Perairan Rawa. *Intek Akuakultur*. 1; 79-89. (2017).
- [10] Slamant, S., M. Marsoedi, A. Mursyid., & D. Arfiati. Konservasi genetik ikan betok (*Anabas testudineus* Bloch 1792) di perairan rawa, Kalimantan Selatan. *Jurnal Penelitian Perikanan Indonesia*. 18; 9-15. (2012).
- [11] Statistik Kementerian Kelautan dan Perikanan (KKP). 2023. Produksi Perikanan.
- [12] <https://statistik.kkp.go.id/home.php> (Diakses tanggal 13 Desember 2022)

Tabel 2. Ikan Papuyu (*Anabas testudineus*) dari lima lokasi di Kalimantan Selatan

Ikan Papuyu (*Anabas testudineus*) dari Kecamatan Candi Laras, Kabupaten Tapin.



Ikan Papuyu (*Anabas testudineus*) dari Desa Lok Baintan, Kecamatan Sungai Tabuk.



Ikan Papuyu (*Anabas testudineus*) dari Kecamatan Danau Panggang, Kabupaten Hulu Sungai Utara.



Ikan Papuyu (*Anabas testudineus*) dari Kecamatan Kurau, Kabupaten Tanah Laut.



Ikan Papuyu (*Anabas testudineus*) dari Desa Batakan, Kabupaten Tanah Laut.



Screening of Phosphate-Solubilizing Fungi from Oil Palm Rhizosphere

Yurnaliza ^{*(1)}, Nunuk Priyani ⁽¹⁾, Irina Levinita ⁽¹⁾

⁽¹⁾*Department Biology, Faculty Mathematics and Natural Sciences, Universitas Sumatera Utara, Jl. Bioteknologi No. 1 Kampus Universitas Sumatera Utara, Medan 20155, Indonesia*

**Corresponding author: yurnaliza@usu.ac.id*

Abstract. The phosphate (P) in soil was fixed by minerals and become available to plant when solubilized by microorganisms such as fungi. The objective of the research was to screening and examining activities of phosphate-solubilizing fungi (PSF) from rhizosphere of oil palm. The location of rhizosphere is in four oil palm plantations near Medan city, North Sumatera Province, Indonesia i. e. plantations in Universities, communities and government owned. Phosphate solubilizing fungi (PSF) was isolated by Pikovskaya medium contain Tricalcium phosphate. Activity of PSF was analyzed qualitatively and quantitatively in solid and liquid Pikovskaya's medium, respectively. The results showed that twenty-two (22) fungal isolates obtained were able to solubilize phosphate that marked by clearing zone near of fungal colony. All of fungi were able to solubilizing phosphate qualitatively with solubilizing index varied from 0,73 until 1,68. Three isolates (TBG06, TBG09 and RIS10) had the highest solubilizing index and also able to solubilizing-phosphate quantitatively in ten days incubation. Measured of pH in fermented medium was showed that activity of phosphate solubilizing was affected by acid condition due to decreasing of pH medium from control in pH 7 to 4.0. *Penicillium* and *Aspergillus* were a two potential Genus of PSF whose produced high concentration of soluble phosphate in fermented culture.

Keywords: oil palm plantation, phosphate solubilizing, fungi, rhizosphere, Pikovskaya medium.

INTRODUCTION

Phosphate (P) is an essential macronutrient next to Nitrogen (N) used for vegetative dan germinative stage of plant growth. Almost more than 95% of soil phosphorus in soil are unavailable to plant [1]. The inorganic phosphorus become insoluble complex when fixed by some cation such as calcium (Ca), aluminum (Al) and iron (Fe) [2]. Chemical reaction some cation caused phosphate become inorganic recalcitrant unavailability for plant root, and the consequences plant became deficiency P element even P are rich in soil.

In the agriculture practice, natural insoluble phosphate such as rock phosphate (RP) is as a common phosphate fertilizer for several purposes. Rock phosphates came from petrified bird or bat droppings. Fertilizing of plant with RP ineffectively if in soil poor of P-solubilizing microorganism [3]. In normal soil pH condition and soil range ranges 5.5 to 6.0 phosphate become soluble but in limited capacity [2]. The role of phosphate solubilizing microorganisms is very important to ensure the availability of dissolved phosphate in various soil conditions.

Phosphate-solubilizing microorganism play as substantial part in accessibility phosphorous in natural soil or inorganic from insoluble state to soluble and available to absorb by root. Several fungi play role as phosphate solubilizing such as Genera *Aspergillus*, *Penicillium*, *Fusarium*, *Mucor* [4]. Munir et al [5] reported 3 species of *Trichoderma* and one species *Pestalotiopsis* displayed the highest solubilization index on plate assay. Solubilizing activity of fungi possess greater abilities to release P from inorganic P, when compared to bacteria. Phosphate solubilization mechanism by P-solubilizing fungi was use some mechanism, and the prominent is through the release of organic acids [5] 6].

To improve phosphate solubility and their availability in soil, P-solubilizing microorganism hardly important in soil mainly to improve plant growth and production of cultivated plant. Phosphate solubilizing fungi in rhizosphere of plant plays an important rule to ensure absorption nutrient to plant [7]. One of cultivated palm in Indonesia that are need more phosphate as fertilizer is oil palm.

Oil palm is an important economic crop plant that produced palm oil for several manufactured industry in Indonesia. Fertilizer management is one of solution to improve production of oil palm fruit bunch. Effectively palm oil production is supported by good management of meeting fertilizer needed such as phosphate. This present research focusses on isolation and screening potential Phosphate Solubilizing Fungi (PSF) from rhizosphere of oil palm. This research is preliminary screening of potency PSF in oil palm rhizosphere for future application in crop management.

MATERIALS AND METHODS

Materials

Soil rhizosphere was collected from four location of oil palm plantation in Medan City, province of North Sumatera, Indonesia i.e. oil palm plant in Universitas Sumatera Utara, Islamic Centre, community-owned Plantation in Tembung and government-owned oil palm plantation in Tanjung Morawa. For amount 1 kg soil from near root of oil palm was collected from each location and then stored in sterile plastic bag and transferred to laboratory.

Isolation phosphate solubilizing fungi

Ten gram of each soil sample was suspended with 90 mL of sterile water in an Erlenmeyer flask and homogenized by shaken at 130 rpm for 30 min. one ml of soil suspension was transferred to 9 ml sterile water to make a serial dilution. One ml suspension from the last series was spread on surface of medium Pikovskaya's agar, made from 10 g glucose, 0.2 g NaCl, 0.5 g yeast extract, 0.2 g KCl, 0.1 g MgSO₄·7 H₂O, 0.5 g (NH₄)₂ SO₄, 0.5 g MnSO₄·7H₂O, 0.5 g FeSO₄·7H₂O, 0.5% tricalcium phosphate, and 18 g agar in 1000 mL of distilled water [8]. Presence of general fungi in soil, each soil suspension also spread on Potato Dextrose agar (PDA) medium. Two type media were incubated at room temperature for 5 days. Total colony fungi in surface medium were calculated. Phosphate solubilizing fungi were identified in Pikovskaya's agar by presence of clear zone around fungal colony. All colony with clear zone or none were calculated and then the fungal with clear zone were cultured to the new medium several times until get the pure culture.

Screening of phosphate solubilizing fungi (PSF)

Active growth mycelium from isolated PSF in 5 mm diameter were sub-cultured to the new the Pikovskaya's agar medium. The experiments were performed in triplicate. The clear zone halo near of the fungal colony and fungal colony diameter were calculated. Index solubilization were determinate by ratio of clear zone diameter and colony diameter.

Characterization of selected isolates

Characterization of PSF were from morphology colony and microscopic. Morphology colony involved color, surface texture, and presence concentric texture. The type hyphae conidia were recorded for characteristic microscopically. Morphological identification of fungi was using the book of Pictorial atlas soil and seed fungi [9].

Phosphate solubilization activity

Selected PSF from initial screening were cultivated on Pikovskaya's liquid medium. Three plug agar media from active growth mycelium were transferred to 25 ml Pikovskaya's liquid sterile in an Erlenmeyer flask. For control was uninoculated. The initial pH medium was ± 6.8 . The all PSF culture medium were incubated at room temperature for 15 days at 150 rpm. All experiments were made in duplicates. Samples were withdrawn every 5 days for a total of 15 days of incubation. From each observation time were calculated pH, soluble phosphate concentration and biomass dry weight. Analysis of P-Solubilizing quantification was conducted by Munir et al [5] procedure.

RESULTS AND DISCUSSION

Isolation and Characterization

Isolation results PSF from 4 locations of rhizosphere oil palm was varied. Total colony of PSF in Pikovskaya agar medium was varied in range 2,0.104 - 3,5.104, abundant in rhizosphere in oil palm di Campus USU (TABLE 1). population PSF in Campus USU when compared with total fungi in the location was 15 %. In Islamic Centre, Tembung and Tanjung Morawa plantation, population PSF respectively 8%, 13 % and 11%.

The total colony with clear zone is less than without clear zone. This is indicated that PSF in soil rhizosphere of oil palm that able to solubilizing fungi is only around 10-20 %.

Table 1. Population of general and Phosphate-solubilizing Fungi from soil rhizosphere of oil palm from 4 locations plantation in Medan City.

Sample location	Total colony in general medium CFU.g-1	Total colony/g soil in Pikovkaya medium (CFU.g-1)	
		With clear zone	Without clear zone
Campus USU	2,2.10 ⁵	3,5.10 ⁴	1,6.10 ⁵
Islamic Centre	2,35.10 ⁵	2,0.10 ⁴	1,3.10 ⁵
Tembung	2,3.10 ⁵	3,0.10 ⁴	1,65.10 ⁵
Tanjung Morawa	2,1.10 ⁵	2,5.10 ⁴	1,15.10 ⁵

Screening of Phosphate Solubilizing Fungi (PSF)

Isolation results was obtained 22 isolates PSF with clear zone when culturing in Pikovkaya agar medium. Solubilizing activity qualitatively from 22 isolates had index solubilization varied in range 0,73 until 1,68. Three isolates with the higher index solubilization value were TBG 06, TBG 09 and RIS 10. Clear zone indicated the Tricalcium phosphate in Pikovskaya agar become soluble. The Genus of phosphate-solubilizing fungi in this research only were characterized in 3 Genus i.e., *Aspergillus* and *Cladosporium*. *Penicillium* and *Aspergillus* are a common phosphate-solubilizing fungi (TABLE 2). To optimization of rock phosphate as P-fertilizer, Mendes et al [10] use *Aspergillus* as a phosphate-solubilizing fungi in culture solid state fermentation with a substrate of sugarcane bagasse. Genus *Penicillium* especially in species *Penicillium bilaiae*, have been producing as commercial product by NovoZymes as a soil inoculant [11].

Table 2. Index phosphate solubilization from selected fungi isolates of oil palm rhizosphere

No	Isolates code	Genus	colony Diameter (mm)	Clear zone diameter (mm)	Solubilization Index
1	USU 01	<i>Penicillium</i>	6,1	8,1	1,32
2	USU 06	<i>Penicillium</i>	7,1	9,2	1,39
3	USU 07	<i>Penicillium</i>	6,9	8,4	1,21
4	USU 12	<i>Aspergillus</i>	5,3	4,8	0,90
5	USU 14	<i>Penicillium</i>	4,9	3,6	0,73
6	USU 15	<i>Penicillium</i>	7,4	10,1	1,36
7	USU 16	<i>Penicillium</i>	5,4	6,7	1,24
8	RIS 02	<i>Cladosporium</i>	4,7	4,5	0,95
9	RIS 03	<i>Penicillium</i>	7,0	9,3	1,32
10	RIS 05	<i>Aspergillus</i>	7,1	9,8	1,38
11	RIS 10	<i>Aspergillus</i>	16,0	23,4	1,48
12	TBG 01	<i>Penicillium</i>	8,0	11,7	1,46
13	TBG 06	<i>Penicillium</i>	5,4	9,1	1,68
14	TBG 07	<i>Penicillium</i>	7,2	9,8	1,36
15	TBG 08	<i>Penicillium</i>	7,4	8,6	1,16
16	TBG 09	<i>Penicillium</i>	8,0	12,0	1,5
17	TM10	<i>Penicillium</i>	2,6	3,2	1,23
18	TM 03	<i>Penicillium</i>	6,7	7,8	1,16
19	TM 04	<i>Penicillium</i>	3,6	4,6	1,27
20	TM 05	<i>Cladosporium</i>	4,0	5,1	1,27
21	TM 06	<i>Penicillium</i>	5,0	5,8	1,16
22	TM 07	<i>Penicillium</i>	11,2	12,8	1,14

Activity of 3 selected isolates to solubilization of phosphate quantitatively showed that solubilization activity of three selected isolates in Pikovskaya liquid medium, indicated that all isolates able to produce the soluble phosphate in high concentration is when incubate in 10 days than 5 or 15 days (TABLE 3). Medium

fermentation become more acid from initial near neutral or 6.8. The pH level is affected phosphorus fixation reaction [12]. The pH level of organic acids produced by fungi greatly influences the level of phosphate solubility. conversely, the solubility level decreases with increasing pH. The $\text{Ca}_3(\text{PO}_4)_2$ can be solubilized by fungi completely when the pH by lower than 7 to 4 [12]. Decreased of pH caused by organic acid that produced by fungi. Elfiati et al [4] reported that two major organic acids produced by phosphate solubilizing fungi are butyric and oxalate acids besides citric, acetic, or malic acid. The different amount fungi produce different types of organic acids, and different types of organic acid had different potency to dissolve Phosphate [13]. The solubilization of the different P sources also depended on the amount of acids production [14]. The soluble phosphate concentration from this research was lower than Elfiati et al [4] [15]. The availability of phosphorus in the soil for plants is not only influenced by the presence of phosphate solubilizing fungi, but is also influenced by more complex factors such as the amount of phosphorus in the soil, the type and amount of soil minerals, soil pH, and other soil factors [3].

TABLE 3. Soluble phosphate concentration, pH and biomass cell of three selected fungi and control at time incubation for 5, 10 and 15 days

No	Isolates code	Incubation (days)	pH	Soluble phosphate concentration (ppm)	Dry cell biomass (mg)
1	Control	5	6,8	0,370	-
		10	6,8	0,423	-
		15	6,8	0,494	-
2	TBG 06	5	5,6	1,151	92
		10	5,2	2,239	159
		15	5,0	1,711	135
3	TBG 09	5	5,0	1,943	76
		10	4,0	2,310	152
		15	5,1	2,224	148
4	RIS 10	5	5,5	1,677	164
		10	5,2	1,801	293
		15	5,6	1,794	178

CONCLUSION

Isolation results of P-solubilizing fungi from 4 rhizosphere of oil palm was obtained 22 isolates of PSF. Solubilizing index from all of PSF isolates varied in range 0,73 until 1,68. Three selected P-solubilizing fungi are from Genus *Aspergillus* (RIS 10) and Genus *Penicillium* (TBG 06 and TBG 09).

REFERENCES

- [1] B. Ikhajiagbe; G.O. Anoliefo O. F. Olise, F. Rackelmann; M. Sommer; I.J. Adekunle, *Notulae Scientia Biologicae* 12(3):500-535 (2020).
- [2] P.D. Johan, O.H. Ahmed, L.Omar, N.A. Hasbullah, *Agronomy* **11**, 2010, 1-25 (2021).
- [3] B.Khoshru, A.F. Nosratabad, D. Mitra, M.Chaithra, Y.R. Danesh, G.Boyno, S. Chattaraj, A.Priyadarshini, S. Andelković, M. Pellegrini, B. E. Guerra-Sierra and S. Sinhaet, *Bacteria* **2**, 98–115 (2023).
- [4] D. Elfiati, A. Susilowati and T. M. Siagian, “Characterization of phosphate solubilizing and cellulolytic fungi isolated from soil under *Eurycoma longifolia* stands”, in *2nd Biennial Conference of Tropical Biodiversity-2021*, IOP Conf. Series: Earth and Environmental Science 886, (IOP Publishing, UK 2021).
- [5] E. Munir, Yurnaliza, A. Lutfia and A. Hartanto, *OnLine Journal of Biological Sciences* **22** (1), (2022).
- [6] N.Vassilev, M. Vassileva, V. Bravo, M. FernandezSerrano and I.Nikolaeva, *Industrial Crops and Products* **26**, 332-336, (2007).
- [7] A. Timofeeva; M. Galyamova; S. Sedykh, *Plants* **11**, 2119 (2022).
- [8] I.Pikovskaya, *Microbiologia*, vol.17,pp.362–370 (1948)
- [9] T.Watanabe (2002) : *Pictorial atlas of soil and seed fungi: morphologies of cultured fungi and key to species*. (CRC Press. Boca Raton) p. 489

-
- [10] G.O.de Mendes, N. M. da Silva, T. C. Anastácio, N. B. Vassilev, J. I. Jr Ribeiro, I. R da Silva, and M. D. Costa, *Microbial biotechnology* **8**(6), 930–939, (2015).
- [11] M. Leggett, J. Cross, G.Hnatowich and G.Holloway, “Challenges in commercializing a phosphate-solubilizing microorganism: *Penicillium bilaiae*, a case history” in *First International Meeting on Microbial Phosphate Solubilization. Developments in Plant and Soil Sciences*, edited by E.Velazquez and C. Rodríguez-Barrueco (Springer, Dordrecht, 2007), pp. 215-222.
- [12] G.O. de Mendes; A.L.M. de Freitas; O..L Pereira; I.R.da Silva; N B Vassilev; M. D. Costa. *Ann Microbiol* **64**, 239–249 (2014).
- [13] D. Elfiati, Delvian, H Hanum, A. Susilowati, H.H. Rachmat, *Biodiversitas* **22**, 3042-3048. (2021).
- [14] Y Zhang; F. S. Chen; X. Q. Wu; F. G. Luan; , L. P. Zhang; X. M. Fang; S. Z. Wan; X. F. Hu; J. R Ye, *PloS one*, **13** (7), . e0199625. (2018).
- [15] F. Elias; D.Woyessa; D. Muleta, *International Journal of Microbiology*, Article ID 5472601. (2016).

Effect of the Methanol Extract of Dragon Scales Leaves (*Drymoglossum piloselloides* [L.] Presl.) on Malondialdehyde and Catalase Level in Blood Serum of Lead-Exposed Albino Rats

Muhammad Syukri Fadil^{*(1)}, Djong Hon Tjong⁽¹⁾, Zozy Aneloi Noli⁽¹⁾, Eti Yerizel⁽²⁾

⁽¹⁾*Biology Departement, FMIPA, Andalas University*

⁽²⁾*Medicine Faculty, Andalas University*

**Corresponding author: msyukrifadil@sci.unand.ac.id*

Abstract. There is massive amount of lead contaminating environment, especially at industrial areas, which is oxidative in its nature. It physiologically effects to the increase of Malondialdehyde and decrease catalase levels in blood serum. To reduce lead poisoning, natural ingredients containing antioxidants are used. Therefore, a study has been carried out regarding antioxidant protective effects of the methanol extract of dragon scales leaves (*Drymoglossum piloselloides* [L.] Presl.) on malondyaldehyde and catalase level in blood serum of albino rats exposed to lead. This study used a completely randomized design (CRD) with 5 treatments and 5 replications each with lead concentration of 0.4% and a dose level of dragon scales leaf extract of 48.51 mg / Kg BW 97.02 mg / Kg BW 194.04 mg / kg BW for 6 weeks. The measurement of Malondialdehyde and catalase levels using spectrophotometric methods. From the treatment results it can be seen that the application of dragon scales leaf methanol extract (*Drymoglossum piloselloides* L.) Presl. can reduce blood serum Malondialdehyde and increase catalase levels in lead-exposed albino rats. From the results it can be concluded that dragon scales leaf extract can be used as an antioxidant.

Keywords: Lead, Malondialdehyde, Catalase, Leaf of Dragon Scales (*Drymoglossum piloselloides* L.) Presl. Antioxidant.

INTRODUCTION

Pb is still widely found in the environment both from metal ore factories resulting from mining, tin smelting, plastics, paints, printing soil and dust as well as in the lead manufacturing plant itself. The use of fuel containing lead, toys, jewelry, ceramics and food contaminated with lead is a source of lead poisoning for urban communities

Exposure to lead into the body can be through the digestive tract and the respiratory route. In the digestive tract can be through food and drink contaminated with lead. Drinking water can be contaminated with Pb which is found in water pipes, solder and water faucets. Pb exposure through the respiratory route can occur in polluted air environments such as on highways and around factories that use lead as raw material. (Zhai, Narbad dan Chen, 2015)

Accumulation of lead in the body is facilitated by the slow excretion of lead in the body. Tissues that are the main target of Pb are hematopoietic tissue, nerves and kidneys. In addition, high accumulation of lead is also found in the liver, lungs, heart, muscles and testes.

The initial mechanism of lead toxicity is the inhibition of the enzyme activity δ -ALAD (delta aminolevulinic acid dehydrogenase). Inhibition of this enzyme increases levels of ALA substrate (aminolevulinic acid). An increase in ALA levels causes the formation of hydrogen peroxide, superoxide radicals and their interactions produce hydroxyl radicals, the most reactive free radicals (Flora, Gupta dan Tiwari. 2012)

In addition, inhibition of δ -ALAD and ferrochelataze enzymes by lead causes a decrease in heme biosynthesis. Hem plays an important role in the biosynthesis of the catalase enzyme which is needed to decompose H₂O₂ (ROS, Reactive Oxygen Species) which are oxidants in the body. Disturbances in the activity of the catalase enzyme will trigger the formation of excess free radicals. An imbalance between natural oxidants and antioxidants in cells will create oxidative stress through two mechanisms, namely formation of reactive oxygen species (ROS) and decreased synthesis of endogenous cell antioxidant enzymes (antioxidant defense systems) such as superoxide dismutase (SOD) and catalase (CAT) (Lamidi dan Akafe, 2017)

Polyunsaturated fatty acids are easily oxidized by free radicals such as H₂O₂. The lipid peroxidation reaction begins with the release of hydrogen atoms from polyunsaturated fatty acids. The lipid radicals formed then react with oxygen to form peroxy radicals. radical chain, when this peroxy radical attracts or removes a hydrogen atom from another fatty acid molecule. The presence of a transition metal such as Fe will initiate

further radical formation. One of the important consequences of lipid peroxidation is the formation of aldehydes, especially malondialdehyde (MDA) (Singh, Kathigesu, Singh and Kaur, 2014)

The effects of free radicals can be reduced by the body by forming antioxidant compounds in the form of enzymatic antioxidants, such as the enzymes superoxide dismutase (SOD), catalase, glutathione peroxidase (GSH-PX), and glutathione reductase (GSH-R). Under conditions of oxidative stress, a way of balancing oxidants and antioxidants is needed by adding exogenous antioxidant compounds from outside the body. These compounds can be obtained from plants that contain antioxidant compounds in the form of phenolic or polyphenolic compounds, namely the flavonoid group such as isoflavones (Kabel, 2014)

To reduce the effects of free radicals on the body due to lead, anti-oxidants from plants are needed. One of the Indonesian plants that has the potential as a source of antioxidants is dragon scales (*Drymoglossum piloselloides* L.). Dragon scales are epiphytic plants from the Polypodiaceae family that live on other trees, firmly attached to large, moist trees such as coconut and mango trees. Dragon scales leaves contain alkaloids, flavonoids, essential oil polyphenols, triterpenes, tannins, sugars and saponins (Azizah, 2016 ; Cahyadi, Bawa dan Sahara, 2014 ; Afrianti, 2015)

Based on research that has been done on dragon scales leaves (*Drymoglossum piloselloides* L) Presl. has antioxidant activity with an IC50 value of 100.76 mcg/ml in the medium category with an IC50 value of 4.299 ppm in the high category and with an IC50 value of 38.29 µg/mL ppm in the high category. However, there has been no report regarding the antioxidant activity test of dragon scales against cell oxidative stress caused by lead. The content of flavonoids and phenolic dragon scales is thought to be able to ward off excessive free radicals and oxidative stress in cells caused mainly by heavy metals. Therefore it is necessary to do research on the antioxidant protective activity of methanol extract of dragon scales leaves (*Drymoglossum piloselloides* L) Presl. on malondialdehyde and catalase levels in blood serum of lead-exposed rats. (Wulandari, Elya, Hanani dan Pawitan, 2013).

MATERIALS AND METHODS

Materials

This study used leaves of dragon scales *Drymoglossum piloselloides*, male rats, rat food in the form of pellets, methanol, ethyl acetate, 100% pure alcohol, 70% to 96% radiant alcohol, distilled water, 2-thiobarbituric acid, glacial acetic acid, sodium hydroxide, malondialdehyde bis and aquabides, Ethanol p.a., Methanol p.a., Toluene, chloroform, Mercury (II) chloride, Bismuth (II) nitrate, Concentrated nitric acid, Iron (III) chloride, Concentrated hydrochloric acid. This research has been submitted to the Faculty of medicine Andalas University Health Research Ethics Commission because this research utilizes experimental animals in its implementation.

Methods

This study used a completely randomized design (CRD) with 5 treatments and 5 replications. The research sample was 25 male Wistar rats (*Rattus norvegicus* L.) selected by simple random technique.

The dose of dragon scales leaf extract of *Drymoglossum piloselloides* given refers to a modified study conducted by Malinda, Fatimawali and Yudistira (2013). For the dose of Pb given, it refers to the research by Mehana, Raheim, Meki and Fazili (2010) and Abdel-Moneim et. al (2014) and the grouping of these treatments is as follows:

- T0 : Group of rats that were only given 0.5% CMC during treatment.
- T1 : Group of rats treated with Pb Acetate concentration of 0.4% during treatment.
- T2 : Group of rats treated with Pb Acetate concentration of 0.4% and dragon scales leaf methanol extract 48.51 mg / Kg BW during the treatment.
- T3 : Group of rats treated with Pb Acetate concentration of 0.4% and dragon scales leaf methanol extract 97.02 mg / Kg BW during the treatment.
- T4 : Group of rats treated with Pb Acetate concentration of 0.4% and dragon scales leaf methanol extract 194.04 mg / Kg BW during the treatment.

MDA measurements were performed on blood using a modified spectrophotometric thiobarbituric acid (TBA) test method. A total of 400 µl of sample was reacted with 200 µl of 20% trichloroacetic acid (TCA) for deproteinization. Then vortex and centrifuge at 5000 rpm for 10 minutes. The supernatant formed was taken and added 400 µL TBA 0.67%. Then the sample was vortexed and incubated in a water heater at 960 C,

10 minutes then removed and cooled to room temperature. Then read the absorption at a wavelength of 530 nm (Jovanović, Nikolić, Kocić, Krstić and Krsmanović, 2012)

Catalase activity was measured in blood, by centrifuging the sample for 10 minutes at 1400 rpm. Then the sample was washed 3 times with 0.9% NaCl and broken down with cold distilled water (1: 4 ratio). The resulting hemolysate is used to determine catalase activity. In this step, a blank solution was also prepared by: Hemolysate (10 µl + 1 ml of distilled water) was added to 500 µl of buffer solution and then the absorbance was measured with a spectrophotometer at $\lambda = 240$ nm (t0). Then the absorbance was measured again after 30 seconds at $\lambda = 240$ nm (t30).

To prepare the test solution, it was carried out in the following way: Hemolysate (10 µl + 1 ml of distilled water) was measured for its absorption with a spectrophotometer at $\lambda = 240$ nm (A0). Then 500 µl H₂O₂ was added and the absorbance was measured again after 30 seconds with a spectrophotometer at $\lambda = 240$ nm (A30). Calculation: Total catalase activity = $\ln(A0 / A30) / \text{sec } t30 - t0$ (Apriana, Supriatin and Rahayu, 2016)

RESULTS AND DISCUSSION

Malondialdehyde levels and catalase levels in blood serum of lead-induced rats and after being treated with methanol extract of dragon scales leaves (*Drymoglossum piloselloides* L) Presl. can be seen in the following table.

Table 1. Average levels of malondialdehyde and catalase enzyme activity in blood serum of lead-induced rats and after being treated with methanol extract of dragon scales leaves (*Drymoglossum piloselloides* L) Presl

Treatment	Malondialdehyde level (nmol/ml)	Catalase activity (Unit/mg)
T0	2,418 ^a	7,600 ^c
T1	3,426 ^c	5,020 ^a
T2	3,118 ^{b,c}	6,420 ^{a,b}
T3	2,992 ^b	7,320 ^b
T4	2,548 ^{a,b}	7,740 ^{b,c}

Note: Numbers not followed by the same lowercase letter show a significant difference in the 5% DNMR test

From the results of the study it was seen that there was an increase in MDA levels in the lead treatment group compared to the T0 group. This may occur because lead is the initiator or induces the oxidation of lipid compounds, especially in unsaturated fatty acids.

Lead that enters the body will form free radicals which produce a reactive oxygen compound (ROS) which has one or more unpaired electrons making it unstable and reactive, namely the hydroxyl radical. These compounds are capable of forming chain reactions by abstracting one hydrogen atom from the cell membrane and forming lipid peroxides. Lipid peroxidation is unstable and easily decomposes to form various compounds, one of which is MDA which has a negative impact on cell membranes (Septian, Thadeus dan widasari, 2012)

There was a decrease in MDA levels in the treatment group given dragon scales leaf methanol extract (*Drymoglossum piloselloides* L) Presl as the treatment dose increased. This shows that the flavonoids present in the methanol extract of dragon scales leaves (*Drymoglossum piloselloides* L) Presl are able to inhibit lipid peroxidation. Flavonoids are able to donate one hydrogen atom from the phenolic hydroxyl (OH) group when reacting with free radicals. Flavonoids play a role in capturing and counteracting free radicals (scavengers) (Lupu and Cremer, 2018)

Flavonoids can reduce the formation of radical compounds and ROS through the decomposition of hydrogen peroxide, quenching active singlet oxygen, and capturing and inactivating radical compounds before they reach the target cell. The latter mechanism is very likely to occur because of the ability of flavonoids to donate hydrogen atoms to radical compounds (El-Latief, 2015)

Another possible mechanism is related to the activity of the enzyme superoxide dismutase (SOD). The increase in SOD activity accelerates the dismutation of superoxide radicals into H₂O₂ and is further eliminated through hepatic catalase activity (Wang, Branicky, Noë and Hakimi, 2018).

The low activity of the catalase enzyme indicates an indirect effect of lead on the catalase enzyme. Pb is known to interfere with heme synthesis, where heme is one of the components in the formation of the catalase enzyme. Disruption of heme synthesis causes a decrease in heme levels in the body, as a result the

antioxidant activity of the catalase enzyme decreases. A decrease in catalase also occurs because lead can replace zinc in the catalase enzyme. This is because the inert free pair of lead binds to the amino acid residues in the enzyme to form different confirmations that change the function of the enzyme. Zinc is the most important cofactor so that the catalase enzyme can act as an antioxidant. So when Zinc is replaced by lead, the catalase enzyme cannot function as an antioxidant. (Flora 2012).

The increase in the catalase enzyme in the treatment after administration of dragon scales leaf methanol extract is closely related to the improvement in heme synthesis which is one of the catalase enzyme formed by flavonoids. The higher the levels of flavonoids, the more of them will chelate lead so that zinc, which is an important cofactor for the catalase enzyme, cannot be replaced by lead.

The mechanism by which flavonoids increase SOD levels in the liver is due to the fact that flavonoids induce antioxidant enzyme genes, then induce antioxidant receptor elements (ARE) and induce DNA to produce antioxidant enzymes. Flavonoids are thought to be able to trigger the expression of antioxidant enzyme genes such as MnSOD, Cu/ZnSOD in the liver so that their activity increases. MnSOD which increases again can cause an increase in H₂O₂ levels as a substrate for catalase. Catalase activity will increase again because catalase will work at a high concentration of H₂O₂ (Zainuri dan Wanandi, 2017).

CONCLUSION

From the results of the research that has been done, it can be concluded that methanol extract of dragon scales leaves ((*Drymoglossum piloselloides* L) Presl can reduce serum malondialdehyde levels in rats exposed to lead. In addition, methanol extract of dragon scales leaves ((*Drymoglossum piloselloides* L) Presl can also increase restore blood serum catalase activity of lead-exposed rats.

ACKNOWLEDGEMENT

Acknowledgments are addressed to the Animal Physiology Laboratory of the Department of Biology FMIPA and the Biochemistry Laboratory of the Medical Faculty of Andalas University for their appreciation and support so that this research can run smoothly.

REFERENCES

- [1] Abdel-Moneim, A. H., A. Meki, A. M. A. S. Gabr, , A. A. Mobasher, M. F. Lutfi, 2014. The Protective Effect of Green Tea Extract against Lead Toxicity in Rats Kidneys. *Asian Journal of Biomedical and Pharmaceutical Sciences*; 4(39) 2014, 30-34.
- [2] Apriana, R., Supriyatin dan Rahayu, S. 2016. Pengaruh Aktivitas Fisik Maksimal dan Ekstrak Daun Bambu Manggong (*Gigantochloa manggong*) Terhadap Aktifitas Katalase Pada Hati Tikus Putih (*Rattus Norvegicus*). *Bioma* 12 (1) 2016; 34-41
- [3] Azizah, N. 2016. Karakter Morfologi Paku Sisik Naga (*Pyrrosia piloselloides*) Berdasarkan Pada Pohon Inang Berbeda. Skripsi Fakultas Sain dan Teknologi UIN Malang.
- [4] Cahyadi, G. A. S. IG. A. G. Bawa dan E. Sahara, 2014. Isolasi dan Identifikasi Senyawa Aktif Anti Bakteri pada Daun Herba Sisik Naga (*Drymoglossum piloselloides* Presl.). *Jurnal kimia* 8 (1), Januari 2014 : 83-90.
- [5] El-Latief, H. M. A. 2015. Protective Effect of Quercetin and or Zinc against Lead Toxicity on Rat Testes. *Global Journal of Pharmacology* 9 (4): 366-376
- [6] Flora, G. D. Gupta, A. Tiwari. 2012. Toxicity of Lead: A Review with recent updates. *Interdicilinary Toxicology*. 2012. Vol.5 (2) 47 – 58
- [7] Jovanović, J.M., Nikolić, R.S., Kocić, G.M., Krstić, N.S., Krsmanović, M.M., 2012, Glutathione protects liver and kidney tissue from cadmium- and lead-provoked lipid peroxidation, *J. Serb. Chem. Soc.* 78(2): 197-207.
- [8] Kabel, A. M., 2014. Free Radicals and Antioxidants: Role of Enzymes and Nutrition. *World Journal of Nutrition and Health*, 2014, Vol. 2, No. 3, 35-38
- [9] Lamidi, I. Y. and I. O. Akafe. 2017. Mitigate Effect of Antioxidants in Lead Toxicity. *MedPub Journals*. Vol 1. No. 1 : 3 ; 1 – 9
- [10] Lupu, A-R and Cremer, L. 2018 Hydroxyl Radical scavenger Activity od Natural SOD. *Romanian Archiever of Microbiology and Immunology*, Vol. 77, Issue 1, pp. 73-83

-
- [11] Malinda, A. F., Fatimawali, Adithya, Y. 2013. Pengaruh Pemberian Ekstrak Etanol Daun Paku Sisik Naga (*Drymoglossum piloselloides* (L) Presl) Terhadap Peroksidasi Lipid Hati Pada Tikus Jantan Galur Wistar Yang Diinduksi CCl₄. *Pharmacol*, Vol. 2 : 72-75
 - [12] Mehana, E.E., A. R. M.A. Meki, K. M. Fazili, 2010. Ameliorated effects of green tea extract on lead induced liver toxicity in rats. *ExpToxicol Pathol* (2010), doi:10.1016/j.etp.2010.09.001
 - [13] Singh, Z., I. P. Kathigesu, P. Singh and R. Kaur, 2014. Use of Malondialdehyde as a Biomarker for Assessing Oxidative Stress in Different Disease Pathologies: a Review. *Iranian J Publ Health*, Vol. 43, Suppl. No.3, Oct 2014, pp: 7-16
 - [14] Septian, A., M. S. Thadeus, dan L. Widasari. 2012. Perubahan Histopatologik Pembuluh Darah Aorta dan Kadar Malondialdehyde (MDA) Pada Mus musculus L. Galur Swiss Derived Akibat Pemberian Minyak Jelantah. *BINA WIDYA*, Volume 24 Nomor 1, Edisi Desember 2012, 6-13
 - [15] Wang, Y., R. Branicky, A. Noë, and S. Hekimi. 2018. Superoxide dismutases: Dual roles in controlling ROS damage and regulating ROS signaling. *J. Cell Biol.* 2018 Vol. 217 No. 6 1915–1928
 - [16] Wulandari, E.T.; Elya, B.; Hanani, E.; Pawitan, J.A. 2013. In Vitro Antioxidant and Cytotoxicity Activity of Extract and Fraction *Pyrrrosia piloselloides* (L) M.G Price. *International Journal of PharmTech Research*. 2013, 5 (1): 119-125.
 - [17] Zainuri, M. dan S. I. Wanandi, 2017. Aktivitas Spesifik Manganese Superoxide Dismutase (MnSOD) dan Katalase Pada Hati Tikus Yang Diinduksi Hipoksia Sistemik: Hubungannya Dengan Kerusakan Oksidatif. *Media Litbang Kesehatan* Volume 22 Nomor 2, Juni Tahun 2012. 87-92
 - [18] Zhai, Q, A. Narbad and W. Chen. 2015. Dietary Strategies for the Treatment of Cadmium and Lead Toxicity. *Nutrients* 2015, 7, 552-571; doi:10.3390/nu7010552

Mercury Binding Antioxidant Enzymes of *Oreochromis Niloticus*: In Silico Evaluation

Noer Komari^{1*}, Eko Suhartono², Suhaili Asmawi³

¹Department of Chemistry, Faculty of Mathematics and Natural Sciences, Lambung Mangkurat University, Jl. A. Yani Km 36 Banjarbaru 70714, Indonesia

²Department of Medicinal Chemistry/Biochemistry, Faculty of Medicine, Lambung Mangkurat University, Jl. A. Yani Km 36 Banjarbaru 70714, Indonesia

³Department of Aquatic Resources Management, Faculty of Fisheries and Marine, Lambung Mangkurat University, Jl. A. Yani Km 36 Banjarbaru 70714, Indonesia.

*Corresponding author: nkomari@ulm.ac.id

Abstract. Mercury enters aquatic organisms through the food chain, with fish being the most susceptible to exposure. Mercury accumulation in fish leads to the production of Reactive Oxygen Species (ROS) and disrupts the activity of antioxidant enzymes in fish, such as Superoxide dismutase (SOD), Catalase (CAT), and Glutathione Peroxidase (GPX). The objective of this study was to investigate the interaction between Hg(II) and amino acid residues within the antioxidant enzyme of tilapia (*Oreochromis niloticus*). This *in silico* study utilizes the protein sequence database available on Uniprot, and enzyme modeling was performed using SWISS-MODEL. The interaction between Hg(II) and the enzyme model was determined through docking simulations using the MIB (Metal Ion-Binding site prediction and docking server). The results revealed that several amino acid residues, including Ala, Arg, Asp, Cys, Glu, Gly, Leu, and Tyr, formed coordination bonds with Hg(II). Additionally, His and Phe residues exhibited cationic interactions with Hg(II). Residues such as Ala, Arg, Asp, Cys, Leu, Gln, Glu, Gly, His, Met, Pro, Ser, Thr, Trp, Tyr, and Val demonstrated hydrophobic interactions with Hg(II). These residues interacting with Hg(II) possess nucleophilic groups, including sulfhydryl groups, carbonyl groups, aromatic groups, and cyclic groups. These findings can serve as valuable insights for exploring interactions between mercury and amino acid residues in other proteins.

Keywords: metal binding, docking molecular, protein modeling, antioxidant enzyme, *Oreochromis niloticus*, Catalase

INTRODUCTION

Mercury (Hg) is a heavy metal toxic in the environment and health [1]. Mercury causes serious effects on cellular, cardiovascular, haematological, pulmonary, renal, immunological, neurological, endocrine, reproductive and embryonic toxicology [2,3]. Mercury enters the aquatic environment from mining, industrial and household activities that flow into the waters through rainwater, river flows and soil erosion [4].

Mercury enters the metabolism of microorganisms and then passes through a fairly long food chain in the end to humans [5]. The main sources of mercury exposure in the human diet are fish and other seafood (>90%) [6]. The toxicity of Hg to fish organs increases with increasing Hg concentration in water. The bioaccumulation of Hg in fish is influenced by the specific physiological and ecological characteristics of different fish species. Mercury exposure can produce teratogenic, neurotoxic, and reproductive toxic effects that can cause damage to cells, tissues, proteins and genes, and ultimately, survival, growth, and behavior of marine fish [7]. Histopathological abnormalities such as hyperplasia, inflammation, and necrosis have been observed in various fish tissues as a result of mercury contamination based on exposure to organic mercury (methylmercury) and inorganic mercury (mercury chloride) through exposure to food or water, respectively, with doses of different duration from mercury[8].

The growth and survival rate of fish decreased with increasing Hg concentration. The number of red blood cells (RBC), hematocrit levels, and hemoglobin levels decreased when compared to controls. The number of white blood cells (WBC) increased in fish given mercury [9,10]. Mercury in the fish body binds to proteins in the form of metal binding proteins. Some of the proteins that bind to Hg include: other isoforms of parvalbumin, ribosomal protein ubiquitin-40S S27a, brain-specific angiogenesis inhibitor-associated protein, and betaine-homocysteine S-methyltransferase. These proteins have a molecular function of binding metal

ions [11]. Several antioxidant enzymes such as Superoxide dismutase (SOD), Catalase (CAT), Glutathione peroxidase (GPX) and Glutathione S-Transferase (GST) also interact with mercury [12]. Characterization of mercury-binding proteins is very important to explain the mechanism of mercury toxicity comprehensively [13]. In fish, Hg affects several enzymes, including antioxidant enzymes, such as GSH and SOD. An increase in blood Hg levels will cause a decrease in the activity of SOD and GSH [14]. Another antioxidant enzyme is Catalase (CAT) which is an important antioxidants for mercury [15]. CAT activity can be significantly inhibited by Hg [16].

Mercury causes oxidative stress and cellular damage in plants, which can be attenuated by the biosynthesis of thiol-rich peptides (biothiols), including glutathione (GSH) [17]. Many studies have linked the interaction of Hg with the -SH (sulfhydryl) group in several enzymes. Fish tissue damage was mainly found in the kidneys, liver and fish exposed to Hg with indications of bleeding, tubular necrosis, vacuolization of epithelial cells, and mononuclear cell infiltration [18]. The mercury ion can bind to the -SH bond and, therefore, decrease the activity and alter the level of sulfur sulfane, the desulfuration product of l-cysteine [19]. Cysteine and related amino acids or their derivatives are important driving forces for the distribution and translocation of Hg in fish [20]. Research on the interaction of heavy metals, including Hg with amino acid residues, has not been carried out extensively. The interaction of Cadmium (Cd) with antioxidant enzymes has been studied previously [21]. This study aims to determine the interaction of mercury with amino acid residues in several antioxidant enzymes of Tilapia (*Oreochromis niloticus*) by molecular docking.

MATERIALS AND METHODS

Protein retrieval and sequence analysis

The protein sequences of SOD, CAT and GPX were obtained from the Uniprot (<https://www.uniprot.org/>) database with the accession codes A0A669EY15, A0A669DUV8, and A0A669D9D1. The physicochemical properties of the protein were determined by the ProtParam web server (<http://web.expasy.org/protparam/>) [22]. The parameters obtained by the ProtParam are molecular weight, theoretical pI, amino acid composition, atomic composition, extinction coefficient, estimated half-life, instability index, aliphatic index, and average hydropathicity (GRAVY).

Prediction of the 3D structure of proteins

Amino acid sequences with FASTA format are submitted to the SWISS-MODEL web server for protein modeling. The 3D structure of the protein will be generated by the homology modeling method. The quality of the model was evaluated with a worksheet on SWISS-MODEL. Model quality is GMQE (Global Model Quality Estimate), QSQE (quaternary structure quality estimation), Sequence Identity, Coverage, Range, Method, Resolution, Oligo-state, Ligands Found, methods by, and Sequence Similarity.

Molecular Docking

Molecular docking using MIB web server: Metal Ion-Binding site prediction and docking server (<http://bioinfo.cmu.edu.tw/MIB/>) [23] to demonstrate the interaction of Hg with amino acid residues of proteins. The result of MIB docking is a template model like the PDB database. The template with the highest score is selected as the model.

Visualization of Model Docking Results

Docking model visualization using NGL viewer software (web GL) in RSCB PDB. The NGL viewer is used to represent the tertiary structure of proteins or peptides and analyze the docking results at the atomic structure level [24].

RESULTS AND DISCUSSION

Physicochemical properties of enzymes

Sequences of the SOD protein of tilapia (*Oreochromis niloticus*) were taken from the Uniprot database with the accession code A0A669EY15. The protein is the Cu-Zn SOD enzyme (EC:1.15.1.1) with the gene code SOD1. The enzyme consists of 153 amino acid residues. The function of the SOD enzyme is to destroy radicals which are normally produced within the cells and which are toxic to biological systems. The SOD enzyme converts superoxide with the help of acid into peroxide and oxygen. SOD enzyme sequence data in FASTA format can be seen in Figure 1.

```
>tr|A0A669EY15|Superoxide dismutase [Cu-Zn],|Oreochromis niloticus
MVLKAVCVLKGTDSTGTVYFEQENESAPVKLTGEIKLTPGEHGFHVHAFGDNTNGCIS
AGPHFNPNYKKNHGGPKDAERHVGDLGNVTAADNVAKIEITDKVITLTGPDSSIQRMTMVIH
EKVDDLKGGNEESLKTGNAGGRLACGVIGITQ
```

Figure 1. Sequence of SOD (A0A669EY15) in FASTA format

The physicochemical properties of the SOD enzymes were determined using the ProtParam web server. The number of amino acids contained in the target protein is 153, the molecular weight is 15918.84. Its isoelectric point (pI) is 5.75. The isoelectric point (pI) is the pH value at which the total charge of the amino acid is zero. If the pH is above the isoelectric point, the protein is negatively charged, while the pH is below the isoelectric point, the protein is positively charged. The estimated half-life of the target protein in mammalian reticulocytes in vitro is 30 hours, in yeast in vivo >20 hours and in *Escherichia coli* in vivo >10 hours. The half-life is the time it takes half of the total protein in the cell to disappear after its synthesis in the cell [25].

The protein instability index (II) is 10.40. This classifies the protein as stable. Instability index (instability index) is a protein measure that provides an estimate of the stability of a protein. A protein whose instability index is less than 40 is predicted to be stable, if it is greater then it is likely to be unstable [26]. Aliphatic index protein is 79.61. The aliphatic index is the relative volume occupied by amino acids such as alanine, valine, isoleucine and leucine, which have aliphatic side chains in their structure and are considered as factors that can increase protein thermostability[25].

GRAVY (grand average of hydropathicity) is a parameter that determines the hydrophobic nature of a protein, if the hydropathic index value of an amino acid is more positive, the more hydrophobic the amino acid is. On the other hand, the more negative the hydropathic index value of the amino acid, the more hydrophilic the amino acid is[27]. The value of the grand average of hydropathicity (GRAVY) protein is -0.313. This value indicates that the protein is hydrophilic. This greatly affects the folding process of an amino acid in the formation of its tertiary structure. Generally, proteins rich in hydrophilic amino acid residues are difficult to fold into a good tertiary (3D) structure. The physicochemical properties of the SOD enzyme are as shown in Table 1.

Table 1. Physicochemical properties of SOD, CAT and GPX

Paramaters	SOD	CAT	GPX
Number of amino acids:	153	538	210
Molecular weight:	15918.84	61177.99	23837.67
Theoretical pI:	5.75	8.06	9.50
Total number of negatively charged residues (Asp + Glu):	19	62	23
Total number of positively charged residues (Arg + Lys):	14	64	33
The estimated half-life is:	30 hours	30 hours	30 hours
The N-terminal of the sequence considered is M (Met).	(mammalian reticulocytes, in vitro). >20 hours (yeast, in vivo).	(mammalian reticulocytes, in vitro). >20 hours (yeast, in vivo).	(mammalian reticulocytes, in vitro). >20 hours (yeast, in vivo).

Paramaters	SOD	CAT	GPX
The instability index (II):	>10 hours (Escherichia coli, in vivo). 10.40 (This classifies the protein as stable)	>10 hours (Escherichia coli, in vivo). 26.80 (This classifies the protein as stable).	>10 hours (Escherichia coli, in vivo). 24.54 (This classifies the protein as stable).
Aliphatic index:	79.61	63.62	77.48
Grand average of hydrophaticity (GRAVY):	-0.313	-0.582	-0.285

Sequences of CAT enzyme *Oreochromis niloticus* were taken from the Uniprot database with code A0A669DUV8. The protein is the enzyme Catalase (EC: 1.11.1.6). The enzyme consists of 538 amino acid residues. CAT enzymes are Occurs in almost all aerobically respiring organisms and serve to protect cells from the toxic effects of hydrogen peroxide. CAT enzymes convert peroxides into water and oxygen. The CAT enzyme sequence in FASTA format can be seen in Figure 2.

```
>tr|A0A669DUV8|Catalase, Oreochromis niloticus
MADNRDKATDQMKTKWKANRGSQKPDVLTGGGHPVGDKLNLTAGPRGPLLVQDVVFTDE
MAHFDRERIPERVVHAKGAGAFGYFEVTHDITRYCKAKVFEHVGGKTTPIAVRFSTVGVFI
VRSWPCDAGESGSADTVRDRPGRFAVKFYTEDGNWDLTGNNTPIFFIRDAMLFPSPVHSQK
RNPQTHMKDPDMVWDFWSLRPESLHQVSFLFSRGLPDGHRHMNGYGSHTFKLVNADGER
VYCKFHYKTNQGIKNLTVEEADRLAATKPDYGIADLFNAIANGNYPSWTFYIQVMTFEQA|
EKFRFNPFDLTKVWSHKEYPLIPVGKLVNRRNPVNYFAEVEQLAFDPSNMPPGIEPSPDK
MLQGRLFSYDPDTHRHLGANYLQIPVNCFFRARVANYQRDGPMMYDNQGGAPNYYPNSE
SAPETQPQFMESKFQVSADVARYNSSDEDNVTQVTRFYTQVLNEEERQRLCQNLAGFLKE
AQLFIQKRMVENLKAVHPDYGNRVETLLKKYNAEAKKDATHVYSRPGAAALAASSKM
```

Figure 2. Sequence of the catalase enzyme from Uniprot with sequence code A0A669DUV8

From the CAT enzyme sequence, the physicochemical properties of the enzyme can be known using the ProtParam web server. The number of amino acids contained in the target protein is 538, the molecular weight is 61,177.99. Its isoelectric point (pI) is 8.06. The isoelectric point (pI) is the pH value at which the total charge of the amino acid is zero. If the pH is above the isoelectric point, the protein is negatively charged, while the pH is below the isoelectric point, the protein is positively charged. The estimated half-life of the target protein in mammalian reticulocytes in vitro is 30 hours, in yeast in vivo >20 hours and in *Escherichia coli* in vivo >10 hours. The half-life is the time it takes for half of the total protein in the cell to disappear after its synthesis in the cell.

The instability index (II) on the target protein was 26.80. This classifies the protein as stable. The instability index is a protein measure that provides an estimate of the stability of a protein. A protein whose instability index is less than 40 is predicted to be stable, if it is greater then it is likely to be unstable [26]. The target protein aliphatic index was 63.62. The aliphatic index is the relative volume occupied by amino acids such as alanine, valine, isoleucine and leucine, which have aliphatic side chains in their structure and are considered as factors that can increase protein thermostability [25].

GRAVY (grand average of hydrophaticity) is a parameter that determines the hydrophobic nature of a protein, if the hydrophatic index value of an amino acid is more positive, the more hydrophobic the amino acid is. On the other hand, the more negative the hydrophatic index value of the amino acid, the more hydrophilic the amino acid is [27]. The value of the grand average of hydrophaticity (GRAVY) of the target protein is -0.582. This value indicates that the target protein is hydrophilic. This greatly affects the folding process of an amino acid in the formation of its tertiary structure. Generally, proteins rich in hydrophilic amino acid residues are difficult to fold into a good tertiary (3D) structure. The physicochemical properties of CAT proteins are shown in Table 1.

Sequences of enzymes GPX of *O. niloticus* fish was taken from Uniprot database with code A0A669D9D1. The protein is the enzyme Protein Glutathione peroxidase (EC:1.11.1.9) *Oreochromis niloticus* (Nile tilapia) (*Tilapia nilotica*) with the gene code GPX8. Enzymes consist of 210 amino acid residues. The

GPX enzyme functions to catalyze peroxidase. GPX converts peroxide to water. The sequence of the CAT enzyme in FASTA format can be seen in Figure 3.

```
>tr|A0A669D9D1| Glutathione peroxidase, Oreochromis niloticus  
  
MEALGGYPTRSSNPKAKKLTVLLSMTVGVGCLFLLQTLVKPRKPKDFYSFEVKDAKGRT  
  
VSLEKYRGKASLVVNVASYCQQTANYKSLQELHRELGTSHFNVLAFPCGQFGDTEGTGS  
  
RDIEAFKASAYGVTFPFFSKIKIMGSEADPAFKFLTDSVKKIPKWNFWKFLVNPEGKVVR  
  
FWRTDEPMESIREEVTAMVREIILKKRVEL
```

Figure 3. Sequence of the GPX enzyme from Uniprot with sequence code A0A669D9D1

From the GPX sequence, the physicochemical properties of the enzyme can be known using the Protparam web server. The number of amino acids contained in the target protein is 210, the molecular weight is 23837.67. Its isoelectric point (pI) is 9.50. The isoelectric point (pI) is the pH value at which the total charge of the amino acid is zero. If the pH is above the isoelectric point, the protein is negatively charged, while the pH is below the isoelectric point, the protein is positively charged. The estimated half-life of the target protein in mammalian reticulocytes in vitro is 30 hours, in yeast in vivo >20 hours and in *Escherichia coli* in vivo >10 hours. The half-life is the time it takes half of the total protein in the cell to disappear after its synthesis in the cell [25].

The instability index (II) on the target protein was 24.54. This classifies the protein as stable. The instability index is a protein measure that provides an estimate of the stability of a protein. A protein whose instability index is less than 40 is predicted to be stable, if it is greater then it is likely to be unstable [26]. The target protein aliphatic index was 77.48. The aliphatic index is the relative volume occupied by amino acids such as alanine, valine, isoleucine and leucine, which have aliphatic side chains in their structure and are considered as factors that can increase protein thermostability [25].

GRAVY (grand average of hydropathicity) is a parameter that determines the hydrophobic nature of a protein, if the hydropathic index value of an amino acid is more positive, the more hydrophobic the amino acid is. On the other hand, the more negative the hydropathic index value of the amino acid, the more hydrophilic the amino acid is (Nelson & Cox, 2007). The value of the grand average of hydropathicity (GRAVY) of the target protein is -0.285. This value indicates that the target protein is hydrophilic. This greatly affects the folding process of an amino acid in the formation of its tertiary structure. Generally, proteins rich in hydrophilic amino acid residues are difficult to fold into a good tertiary (3D) structure. The physicochemical properties of SOD enzyme is as shown in Table 1.

Antioxidant enzymes modeling

The SOD enzyme was modeled with the SWISS-MODEL web server based on the sequence with FASTA format as shown in Figure 1. The model selected from the modeling results was a model with template code 5yto.1.A. The model is Superoxide dismutase [Cu-Zn] Crystal Structure Human Superoxide Dismutase I (hSOD1) in complex with naphthalene-catechol related compounds. The model with code 5yto.1.A is considered the best model with parameters GMQE= 0.88, QSQE= 0.95, Seq Id= 70.59, Coverage= 1.00, Range= 1-153, Method= X-ray Resolution= 1.90 Oligo-state= homo -dimer, Ligand Found= 2 x ZN, by= BLAST, Equation Seq= 0.52. The model image is shown in Figure 4A.

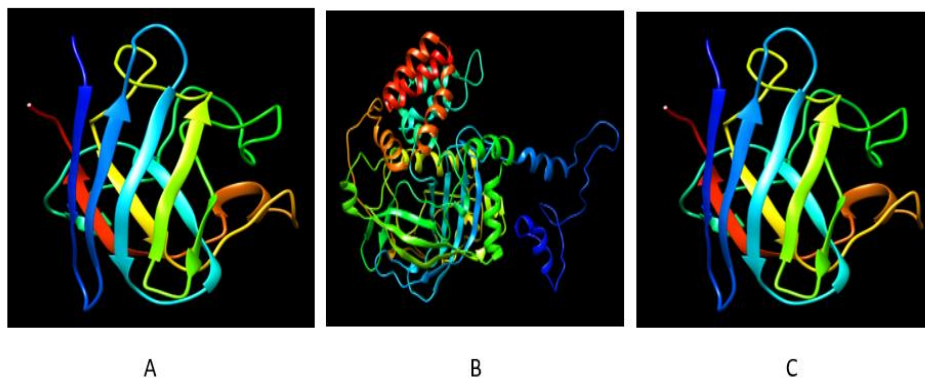


Figure 4. Enzyme Model of *Oreochromis niloticus*, A=SOD, B=CAT, C=GPX

The CAT enzyme was modeled with the SWISS-MODEL web server based on the sequence with the FASTA format as shown in Figure 2. The model selected from the modeling results was a model with the template code 7cat.1.A. The model is CATALASE, The NADPH binding site on beef liver catalase. Model with code 7cat.1.A. considered as the best model with parameters GMQE= 0.92, QSQE= 0.89, Seq Id= 83.79, Coverage= 0.94, Range= 2-518, Method= X-ray Resolution= 2.50 Oligo-state= homo-tetramer, Ligands Found= 4 x HEM, 4 x NDP , by= BLAST, Seq Similarity= 0.57. The model is shown in Figure 4B. Furthermore, the GPX enzyme was modeled with the SWISS-MODEL web server based on the sequence with FASTA format as shown in Figure 3. The model selected from the modeling results was a model with 3KIJ template code. The model is Crystal structure of the human PDI-peroxidase. The model with code 3KIJ is considered the best model with parameters GMQE= 0.73, QSQE= 0.00, Sequence Identity= 65.06, Coverage= 0.79, Range= 42-207, Method= X-ray, Resolution= 1.80, Oligo-state= monomer, Ligands Found= No, by methods= HHblits, Sequenece Similarity= 0.51. The model image is shown in Figure 4C.

Interaction of Mercury with SOD

The interaction of Hg(II) and SOD enzyme modeled with SWISS-MODEL was analyzed by MIB. Table 2 shows the template for the MIB analysis docking results. Binding residues are approximate amino acid residues that interact strongly with Hg. The template is a 3D structure model contained in the PDB RCSB that has the same or similar to the docking results. The highest score shows the best template from the results of docking Hg with SOD. The 1DKN template was chosen because it has the highest score, 1.044.

Table 2. Docking results Hg(II) with MIB

Superoxide dismutase				Catalase			Glutathione peroxidase		
No	Residues	Template (PDB ID)	Score	Residu	Template (PDB ID)	score	Residues	Template (PDB ID)	Score
1	81H, 85L	1DKN	1.044	467R, 471C	3KBK	1.805	145G, 146S	2J4F	1.490
2	58C, 143R	1DLQ	0.948	378G, 379A	2J4F	1.255	80C, 109C, 116E	3ZNB	0.840
3	143R, 146C, 148V	4M1M	0.914	126C, 137V	4M1M	1.070	79Y, 80C	1HT3	0.771
4	7C, 147G	2PGV	0.908	94Y, 95C	1HT3	1.068	80C, 114D	1SDN	0.672

Visualization of docking results with NGL viewer on RCSB PDB with code (1DKN) is presented in Figure 5. Visual analysis of the interaction of Hg(II) with the SOD enzyme shows 5 types of interaction models. There are 2 types of bonds in the Hg and SOD interactions, namely metal coordination covalent bonds and hydrophobic interactions. The interaction in model 1 (Figure 5A) shows the Tyrosine (Y) residue with Hg(II) forming a metal coordination covalent bond; histidine (H) and glycine (G) form hydrophobic

interactions with Hg(II). Tyrosine has a carbonyl group on the carboxylate which is capable of forming a metal coordination covalent bond with Hg(II).

The interaction in model 2 (Figure 5B) shows that D (aspartic acid) and E (glutamic acid) form metal coordination covalent bonds with Hg(II); Other R (arginine), T (Threonine), H (histidine) and D (aspartic acid) form hydrophobic interactions with Hg(II). The two carbonyl groups, C=O in both aspartic acid and glutamic acid carboxylic acids are capable of forming metal coordination covalent bonds with Hg(II). The interaction in model 2 shows that D (aspartic acid) and E (glutamic acid) form metal coordination covalent bonds with Hg(II); Other R (arginine), T (Threonine), H (histidine) and D (aspartic acid) form hydrophobic interactions with Hg(II). The two carbonyl groups, C=O in both aspartic acid and glutamic acid carboxylic acids are capable of forming metal coordination covalent bonds with Hg(II).

Model 3 (Fig. 5C) shows the Leucine (L) residue forming a metal coordination covalent bond with Hg(II); proline (P), Glutamine (Q), Histidine (H), and Threonine (T) form hydrophobic interactions with Hg(II). Leucine (L) has a carbonyl group, C=O on the carboxylate which forms a metal coordination covalent bond with Hg(II). Model 4 (Fig. 3D) shows alanine (A) residues forming a metal coordination covalent bond with Hg(II); proline (P), Leucine (L), threonine (T), and serine (S) form hydrophobic interactions with Hg(II). Alanine (A) has a carbonyl group, C=O on the carboxylate which forms a metal coordination covalent bond with Hg(II). Model 5 (Figure 4 E) shows Hg(II) only forming hydrophobic interactions with residues Valina (V), leucine (L), threonine (T), alanine (A), and serine (S).

The interaction of amino acid residues with Hg(II) is dominated by electron-rich groups, including -SH and -C=O groups. Almost all amino acid residues that have -SH and C=O groups can bind to Hg(II) forming metal coordination covalent bonds and hydrophobic interactions. Soft metal ions of class B such as Hg(II) (according to the Lewis acid classification) preferentially form covalent bonds with oxygen, sulphur or nitrogen donors and form rather stable complexes[28]. Hg's affinity for thiol groups may also underlie the pathway of Hg toxicity. Mercury-induced oxidative stress can occur from Cys-Hg binding and inhibition of Mn-SOD (Cys196) activity [18].

In addition, the phi cation bond was also found in the aromatic residue group of phenylalanine with Hg(II). The structural description of the interaction of Hg(II) with amino acid residues will be useful for designing mercury chelators that are more effective as additional binders and hydrophobic protecting groups in binding mercury(II) [29]. Mercury can interact with various amino acid residual in all three types of hexokinase enzyme. Most of amino acid residual that is contained in hexokinase enzyme form a coordinate covalent bond with and mercury. The metal bond makes heavy metal trapped in hexokinase enzyme, hence the metal is hard to be unattached [30].

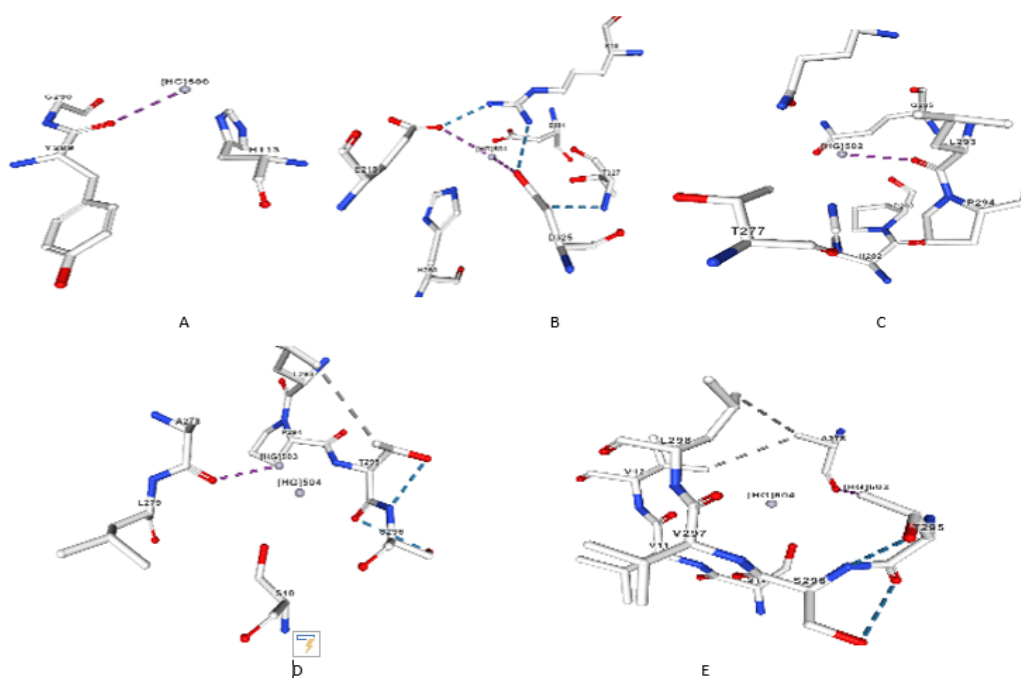


Figure 5. Visualization of docking results Hg(II) with SOD; A = model 1, B= model 2, C=model 3, D= model 4 and E = model 5

Interaction of Hg and CAT

The interaction of Hg and CAT enzyme modeled with SWISS-MODEL was analyzed by MIB. Table 2 shows the template for the MIB analysis results. Binding residues are approximate amino acid residues that interact strongly with Hg. The template is a 3D structure model contained in the PDB RCSB that has the same or similar to the docking results. The highest score shows the best template from the results of docking Hg with CAT. The 3KBK template was chosen because it has the highest score, 1.805.

Visual analysis of Hg docking results with CAT enzymes with 3kbk template showed 4 types of models (Figure 6). There are 3 types of bonds in the interaction of Hg (II) and CAT enzymes, namely metal coordination covalent bonds, phi cation interactions and hydrophobic interactions. The interaction in model 1 (Figure 6A) shows that the cysteine (C) residue forms a metal coordination covalent bond with Hg(II); tyrosine (T), alanine (A) and tyrosine (Y) form hydrophobic interactions with Hg(II). The sulfhydryl group on cysteine forms a metal coordination covalent bond with Hg(II).

The interaction in model 2 (Figure 6B) shows that E (glutamic acid), R (arginine) and C (cysteine) form metal coordination covalent bonds with Hg(II); Other L (leucine), E (glutamic acid), V (valine), and W (tryptophan) form hydrophobic interactions with Hg(II). The carbonyl group, C=O in the carboxylic acid aspartic acid and glutamic acid forms a metal coordination covalent bond with Hg(II); the -SH group on cysteine forms a hydrophobic interaction with Hg(II). Model 3 (Figure 6C) shows the Cys(C) residue forming a metal coordination covalent bond with Hg(II); His (H) forms a phi cation interaction and Asn (N), Tyr (Y) forms a hydrophobic interaction with Hg(II). Cys has an -SH group which forms a metal coordination covalent bond with Hg(II). The cyclic group on histidine b forms the interaction of the phi cation with Hg(II). Model 4 (Figure 6D) shows the Cys(C) residue forming a metal coordination covalent bond with Hg(II); Arg (R), Leucine (L), Glu (E), and Glycine (G) form hydrophobic interactions with Hg(II). Cys (C) has an SH group that forms a metal-covalent covalent bond with Hg(II). Hg is generally bound to cysteine residues, because the Hg-cysteine complex has a high stability constant. The formation of the Hg-cysteine complex is caused by Hg binds to a free thiol group which available in cysteine. Mercury bonds in the group thiol residue of cysteine causes the cysteine residue to malfunction in proteins, where the thiol group of cysteine is the active group of most enzyme. Mercury causes enzymes not active because the active side is not working again [31].

The template in Table 2 shows several amino acid residues that interact with Hg, including: Arg (R), Cys (C), Gly (G), Ala (A), Val (V), Tyr (Y), and His (H), while in PDB the amino acid residues that interact with Hg are more varied, namely: Cys, Thr T, Ala, Tyr (Y), GluE, Arg R, Leu L, Val V, Trp W, His H, Asn N, Gly G.

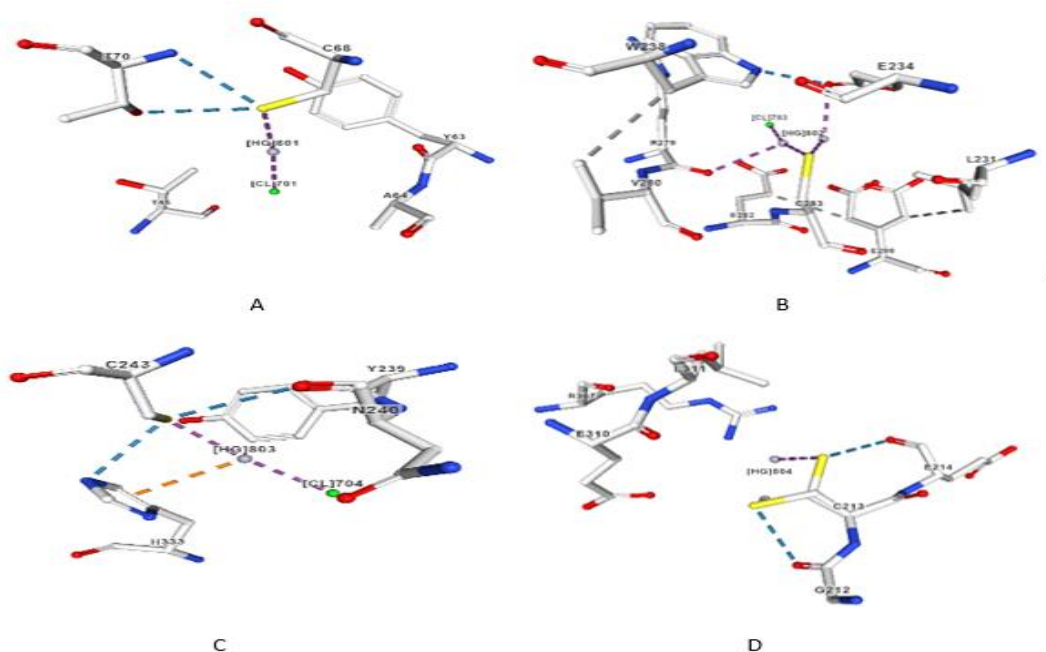


Figure 6. Visualization of docking results Hg(II) with CAT; A = model 1, B= model 2, C=model 3, and D= model 4

Interaction of Hg with GPX.

The interaction of Hg(II) and GPX enzyme modeled with SWISS-MODEL was analyzed by MIB. Table 2 shows the template for the MIB analysis results. Binding residues are approximate amino acid residues that interact strongly with Hg. The template is a 3D structure model contained in the PDB RCSB that has the same or similar to the docking results. The highest score shows the best template from the results of docking Hg with GPX. The 2J4F template was chosen because it has the highest score, 1.490.

Visual analysis of Hg(II) docking results with CAT shows 2 types of interaction models (Figure 7). There are 2 types of bonds in the Hg and CAT interactions, namely metal coordination covalent bonds and hydrophobic interactions. Figure 7A is an interaction model 1, showing that the Glu(E) residue forms a metal coordination covalent bond with Hg(II); His (H) and tyrosine (T), form hydrophobic interactions with Hg(II). Glutamic acid (E) has a C=O carbonyl group on the carboxylate which can form a metal coordination bond with Hg(II). Figure 7B is an interaction model 2, showing that two residues Serine (S) and C (cysteine) form a metal coordination covalent bond with Hg(II); W (tryptophan), N (asparagine), V (Valine), Proline (P), G (glycine) and Alanine (A) form hydrophobic interactions with Hg(II). The carbonyl group C=O on the carboxylate of serine forms a metal coordination covalent bond with Hg(II); In cysteine the -SH group binds to Hg(II). Mercury(II) can form complexes with oxygen, nitrogen or sulfur donor ligands [32]. Cysteine (Cys) is considered to be the main target of Hg, and relates its sensitivity to thiol groups and cell damage. Binding of Hg to Cys mediates some of the toxic effects of Hg, particularly its inhibitory effect on enzymes and other proteins containing free Cys residues [18]. The active site of the enzyme contains a sulfhydryl group at cysteine residues will bind covalent to metal. Affinity high in mercury in the sulfhydryl group at the catalytic site of the enzyme is the motif major commonly known in enzyme inactivation [31].

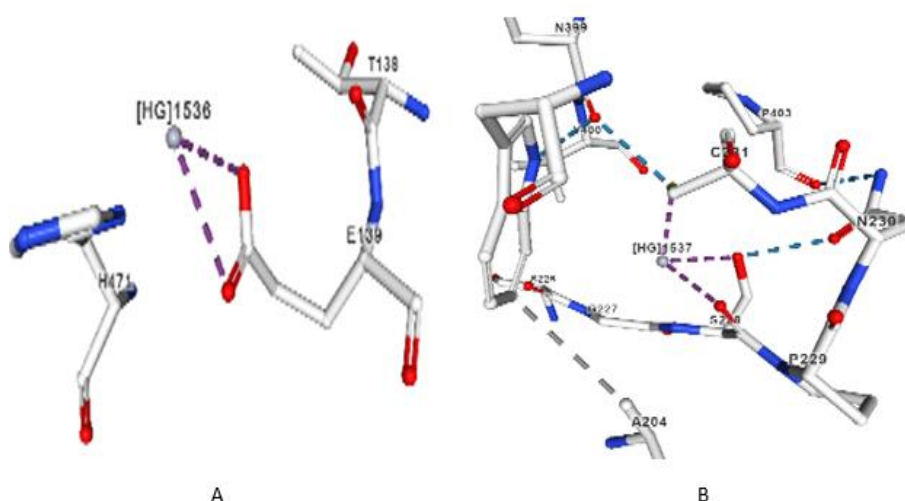


Figure 7. Visualization of docking results Hg(II) with GPX; A = model 1, B= model 2,

CONCLUSION

The *in silico* study of the interaction between mercury and antioxidant enzymes shows the presence of a strong metal binding between mercury and enzyme residues. Hg(II) interacts with amino acid residues found in the antioxidant enzymes of tilapia (*Oreochromis niloticus*) within electron-rich amino acid groups, known for their nucleophilic characteristics. These groups encompass -SH, -C=O, aromatic rings, and cyclic rings. Amino acid residues that bind to Hg(II) include: Cys, Thr, Ala, Tyr, Glu, Arg, Leu, Val, Trp, His, Asn, Gly, Gln. The results of this study can be used for further examining the interaction of mercury with other proteins and studying the mechanisms of mercury toxicity in organisms.

ACKNOWLEDGEMENT

Thank you to Mr. Abdul Gafur, the Dean of the Faculty of Mathematics and Natural Sciences who has facilitated the use of the bioinformatics laboratory.

REFERENCES

- [1] Leong, Y. K. & Chang, J. S. Bioremediation of heavy metals using microalgae: Recent advances and mechanisms. *Bioresource Technology* (2020). doi:10.1016/j.biortech.2020.122886
- [2] Ye, B. J. *et al.* Evaluation of mercury exposure level, clinical diagnosis and treatment for mercury intoxication. *Annals of Occupational and Environmental Medicine* **28**, (2016).
- [3] Genchi, G., Sinicropi, M. S., Carocci, A., Lauria, G. & Catalano, A. Mercury exposure and heart diseases. *International Journal of Environmental Research and Public Health* **14**, (2017).
- [4] Ibrahim, T. A. & Aris, M. Toksisitas Merkuri (Hg) pada struktur jaringan ikan. *e-Journal Budid. Perair.* **9**, (2020).
- [5] Rice, K. M., Walker, E. M., Wu, M., Gillette, C. & Blough, E. R. Environmental mercury and its toxic effects. *Journal of Preventive Medicine and Public Health* **47**, (2014).
- [6] Grieb, T. M., Fisher, N. S., Karimi, R. & Levin, L. An assessment of temporal trends in mercury concentrations in fish. *Ecotoxicology* **29**, (2020).
- [7] Zheng, N. *et al.* The Toxicological Effects of Mercury Exposure in Marine Fish. *Bulletin of Environmental Contamination and Toxicology* **102**, (2019).
- [8] Zulkipli, S. Z. *et al.* A review of mercury pathological effects on organs specific of fishes. *Environmental Pollutants and Bioavailability* **33**, (2021).
- [9] Nirmala, K., Hastuti, Y. P. & Yuniar, V. Toxicity of mercury (Hg) on survival and growth rate, hemato- and histopathological parameters of *Oreochromis niloticus*. *J. Akuakultur Indones.* **11**, (2013).
- [10] Kaoud, H. A., EL-Dahshan, A. R. & Ahmed, Q. Mercury toxicity in aquatic environment: histopathology, hematological and enzymatic alterations in *Oreochromis niloticus*; methods of amelioration. *Glob. J. Multidiscip. Appl. Sci.* **2**, 59–65 (2014).
- [11] Vieira, J. C. S. *et al.* Identification of protein biomarkers of mercury toxicity in fish. *Environ. Chem. Lett.* **15**, (2017).
- [12] Bittarello, A. C. *et al.* Metalloproteomic approach of mercury-binding proteins in liver and kidney tissues of *Plagioscion squamosissimus* (corvina) and *Colossoma macropomum* (tambaqui) from Amazon region: Possible identification of mercury contamination biomarkers. *Sci. Total Environ.* **711**, (2020).
- [13] Li, Y. *et al.* Characterization of mercury-binding proteins in human neuroblastoma SK-N-SH cells with immobilized metal affinity chromatography. *Talanta* **178**, (2018).
- [14] Henry, K. A., Cristol, D. A., Varian-Ramos, C. W. & Bradley, E. L. Oxidative stress in songbirds exposed to dietary methylmercury. *Ecotoxicology* **24**, (2015).
- [15] Wigfield, D. C. & Perkins, S. L. Oxidation of mercury by catalase and peroxidase in homogeneous solution. *J. Appl. Toxicol.* **3**, (1983).
- [16] Chen, L., Zhang, J., Zhu, Y. & Zhang, Y. Molecular interaction of inorganic mercury(II) with catalase: A spectroscopic study in combination with molecular docking. *RSC Adv.* **5**, (2015).
- [17] Sobrino-Plata, J. *et al.* Sulphur and biothiol metabolism determine toxicity responses and fate of mercury in *Arabidopsis*. *Environ. Exp. Bot.* **182**, 104302 (2021).
- [18] Ajsuvakova, O. P. *et al.* Sulfhydryl groups as targets of mercury toxicity. *Coordination Chemistry Reviews* **417**, (2020).
- [19] Sura, P., Bronowicka-Adamska, P., Furtak, E. & Wróbel, M. Effect of mercury ions on cysteine metabolism in *Xenopus laevis* tissues. *Comp. Biochem. Physiol. Part C Toxicol. Pharmacol.* **154**, 180–186 (2011).
- [20] Man, Y. *et al.* Primary amino acids affect the distribution of methylmercury rather than inorganic mercury among tissues of two farmed-raised fish species. *Chemosphere* **225**, (2019).
- [21] Komari, N. & Suhartono, E. Cadmium binding to antioxidant enzymes: in silico study. *IOP Conf. Ser. Mater. Sci. Eng.* **980**, 012038 (2021).
- [22] Gasteiger, E. *et al.* Protein Identification and Analysis Tools on the ExPASy Server. in *The Proteomics Protocols Handbook* (2005). doi:10.1385/1-59259-890-0:571
- [23] Lin, Y. F. *et al.* MIB: Metal Ion-Binding Site Prediction and Docking Server. *J. Chem. Inf. Model.* **56**, (2016).
- [24] Rose, A. S. *et al.* NGL viewer: Web-based molecular graphics for large complexes. *Bioinformatics* **34**, 3755–3758 (2018).
- [25] Walker, J. M. *Protein Protocols Handbook, The. Protein Protocols Handbook, The* (2002). doi:10.1385/1592591698
- [26] Guruprasad, K., Reddy, B. V. B. & Pandit, M. W. Correlation between stability of a protein and its

- dipeptide composition: A novel approach for predicting in vivo stability of a protein from its primary sequence. *Protein Eng. Des. Sel.* (1990). doi:10.1093/protein/4.2.155
- [27] Nelson, D. L. & Cox, M. M. *Lehninger Principles of Biochemistry* 7th. *W.H. Free. Co.* **2**, (2017).
- [28] Nieboer, E. & Richardson, D. H. S. The replacement of the nondescript term 'heavy metals' by a biologically and chemically significant classification of metal ions. *Environ. Pollut. Ser. B, Chem. Phys.* **1**, 3–26 (1980).
- [29] Ngu-Schwemlein, M., Merle, J., Meeker, W., Risdon-Langdon, K. & Nixon, T. Evaluating the involvement of tryptophan on thiolated peptide-mercury(II) complexes: Cation- π interactions. *Inorganica Chim. Acta* **506**, 119552 (2020).
- [30] Pratidina, E. A., Suhartono, E. & Setiawan, B. IMPACT OF HEAVY METALS ON HEXOKINASE ISOFORMS: AN IN SILICO STUDY. *Berk. Kedokt.* **18**, (2022).
- [31] Suhartono, E., Komari, N. & Siahaan, S. C. P. T. Interaksi merkuri dan kadmium terhadap enzim kunci pada glikolisis in siliko. *J. Ilm. Kedokt. Wijaya Kusuma* **10(2)** **2071**, (2021).
- [32] Sajidu, S. M. I., Persson, I., Masamba, W. R. L. & Henry, E. M. T. Mechanisms for biosorption of chromium(III), copper(II) and mercury(II) using water extracts of *Moringa oleifera* seed powder. *African J. Biotechnol.* **7**, (2008).

Calotropis gigantea Derived Carbon Dots For Breast Cancer Bioimaging

Musbahu Adam Ahmad¹, Sri Sumarsih¹, Siti Febtria Asrini Sugito¹, Aswandi Wibrianto²,
Mochamad Zakki Fahmi^{*1}

¹Department of chemistry, FST, 60115, Universitas Airlangga.

²Department of Chemical Engineering, National Taiwan University of Science and Technology,
Taipei 106335, Taiwan, Republic of China.

*Corresponding author: m.zakki.fahmi@fst.unair.ac.id

Abstract: We aim to study the synthesis and application of carbon dots derived from dried latex of *C. gigantea* (G-CDs) for cancer diagnosis. Previously, we established that carbon dots prepared from biologically active sources often tend to possess similar to better activity compared to the starting material [19]. *C. gigantea* is a shrub whose latex is rich in constituents with anticancer biological application. Thus, this could be an invaluable source for carbon dots with application in cancer diagnosis. The carbon dots were synthesized by pyrolysis method. Then the CDs were further characterized by UV-vis and FTIR spectroscopy; PL spectroscopy and DLS. Furthermore, confocal microscopy was used for imaging breast cancer cell lines (MDA-MB-231). G-CDs showed excellent stability at different conditions of pH, temperature and ionic strength. DLS indicated that the CDs have a size around 8nm. PL confirmed fluorescence property with QY of 21%. FTIR indicates the presence of various functional groups that include hydroxyl (3250 cm⁻¹) and carbonyl groups (1650 cm⁻¹). CLSM showed that CDs successfully entered into MDA-MB-231 cell lines and fluorescence as they were excited with laser light. In summary, latex sourced fluorescent carbon dots with potential application in breast cancer diagnosis were synthesized.

Keyword: Carbon dots, fluorescence, Quantum Yield, Confocal Microscopy, Bioimaging, Diagnosis

INTRODUCTION

Cancer is still among the top diseases with continuous increase of cases and mortality globally. As of 2022, cancer was estimated to reach about 2 million new cases with accompanying deaths of over 600 thousand in The United States[1]. Furthermore, statistics from Indonesia also indicate the huge impact of cancer on the well being of people of Indonesia. Globocan cancer data (2020) reported an estimate of about 59 percent associated deaths due to cancer[2]. Instances from these two different regions have made the necessity to continue finding alternatives for combating cancer; especially when underdeveloped countries are put into considerations. Since the causes of cancer remain ambiguous, one of the best ways to prevent the threats posed by cancer is by detecting it as early as possible. This would be feasible if cheap, easy and effective detection methods are employed.

Carbon dots, on the other hand, are nanomaterials that are based mainly on carbon core structure. They are tiny (below 10nm) roughly spherical nanomaterials that have taken the interest of many researchers due to their versatile properties. Specifically, carbon dots have excellent biocompatibility and fluorescence behavior. This property hinted at the potential application of carbon dots for bioimaging. In addition, their amenable surface makes targeting cells easier. A number of researches have been reported which uses fluorescent carbon dots for application in bioimaging cancer cells. For instance, a recent review by Masoudi Asal group [3] discussed different types of carbon nanomaterials for application in diagnosis and therapy. They unambiguously stated the potency of CDs in diagnosing different kinds of cancers. Furthermore, fluorescent CDs derived from folic acid were used for imaging breast cancer cell lines[4]. Similarly, CDs derived from spices (red chili, cinnamon, black pepper, and turmeric) were used for fluorescent imaging of human glioblastoma (LN-229) cell lines[5]. Karakoçak *et al.*[6] synthesized nitrogen-doped CDs conjugated with hyaluronic acid for targeted imaging of CD44 receptor-expressing breast cancer cell lines. Recently, *Azadirachta indica* (Neem) leaves were utilized for the synthesis of CDs for multipurpose applications. Aside from being potentially antioxidant and antimicrobial agents, the CDs conjugated with folic acid were able to target MCF-7 breast cancer cell lines for bioimaging application. In summary, fluorescent CDs could be synthesized from a wide range of sources. Synthesis of CDs from sustainable sources with valuable application in cancer therapy and diagnosis could assuage some of the challenges faced during combating cancer diseases.

Calotropis gigantea is one of the two species of *Calotropis* genus. It is commonly known as Yercum fibre in English and “biduri” or “widuri” in Indonesian. *C. gigantea* is an invasive shrub that grows in arid and semi-arid regions. Various parts of this plant include leaves, flowers, stem, root, fruit and latex. One or more parts of this plant have been used traditionally for treating various ailments. Diseases like leprosy, cancer, piles and asthma were treated using this plant[7]. The latex of *C. gigantea* is white thick liquid that comes out when a part of the plant is cut. It is rich in multiple secondary metabolites with cytotoxicity towards cancer cells.[8] Therefore, being a resilient invasive species that could grow in different harsh conditions makes its utilization for various applications sustainable. To date, few researches were conducted to synthesize CDs from latex of this plant for biomedical applications. Murugesan and co-researchers developed highly biocompatible CDs from the plant latex for metal sensing and radical scavenging activity[9]. Furthermore, latex derived CDs were used for staining fungi, bacteria and plants[10]. However, to the best of our knowledge, application of latex derived CDs for potential application in cancer cell imaging has not yet been studied. Therefore, in this research, we aim to synthesize CDs for cancer cell bioimaging. The synthesized CDs would be characterized by techniques consisting of Ultra-violet-visible (UV-vis) spectroscopy; photoluminescence spectroscopy; Fourier transform infrared spectroscopy (FTIR), and Dynamic Light Scanning (DLS). Finally, the fluorescent CDs would be used for imaging MDA-MB-231 breast cancer cell lines.

MATERIALS AND METHOD

Materials

Ethanol (99.5%; C₂H₅OH), and 4',6-diamidino-2'-phenylindole dihydrochloride (DAPI) were purchased from Sigma-Aldrich, USA. Hydrochloric acid (37%; HCl) and sodium hydroxide (NaOH, 98%) were purchased from Merck, Germany. Dulbecco's modified Eagle's cell medium and phosphate buffered saline (PBS) were purchased from GIBCO Thermo Scientific, USA. All chemicals were used directly without further purification.

Sample Collection and Preparation

The latex of *C. gigantea* was directly collected directly from the plant at a land located opposite to faculty of fisheries and Aquaculture Universitas Airlangga Kampus C. The latex was dried at normal room conditions and ground to powder.

Synthesis of Carbon dots from the Dried Latex

CDs from *C. gigantea* latex (G-CDs) were prepared by method of pyrolysis described as follows. 1g of the powdered latex was put in a vial and heated in a muffle furnace for 2 hours at 300°C. Then the particles were then dissolved with a few quantities of NaOH followed by ultrasonication to ensure proper dissolution. The solution was then centrifuged and the supernatant was further filtered through 0.2µm filter membrane. Finally the prepared G-CDs were dried to obtain solid CDs.

Characterizations

Hydrodynamic diameter of the CDs was measured by using DLS Nanoparticle Size Analyzer (Biobase, China). PL spectra of G-CDs were measured using the BK-F96Pro Fluorescence Spectrophotometer (Biobase, China) equipped with a xenon lamp. The fluorescence of G-CDs solution was recorded at various excitation wavelengths. The QY was calculated using the standard reference method with equation 1, which uses the integrated area of the PL curve[11].

$$\varphi_{\text{sample}} = \varphi_{\text{reference}} \left(\frac{\text{Abs}_{\text{reference}}}{\text{Area}_{\text{reference}}} \right) \left(\frac{\text{Area}_{\text{sample}}}{\text{Abs}_{\text{sample}}} \right) \quad (1)$$

where φ_{sample} and $\varphi_{\text{reference}}$ are QY of the sample and reference respectively. Whereas $\text{Abs}_{\text{reference}}$; $\text{Abs}_{\text{sample}}$ as well as $\text{Area}_{\text{reference}}$ and $\text{Area}_{\text{sample}}$ are their absorbances and integrated areas of the PL curve respectively.

Fourier transformed infrared (FTIR) spectrum was observed with an FTIR Spectrophotometer (Shimadzu, Japan). 1mg of the sample was made into pellets using 100mg KBr followed by FTIR analysis. UV-Visible Spectral absorption was recorded using Shimadzu UV-1800 (Shimadzu Corporation,

Japan). Colloidal stability of the CDs was tested by subjecting them to different pH conditions (4–12), different temperature ranges (30–90°C), and varied concentrations of NaCl (0 - 0.5 M).

Confocal Laser Scanning Microscopy

Breast cancer cell lines (MDA-MB-231) were cultivated in a six-well plate for imaging. The cell lines (MDA-MB-231) were grown in DMEM medium at a density of 1.0×10^6 and incubated for 24hr. Next, the medium was washed with PBS followed by introducing a new medium containing 200 ppm of G-CDs and incubated for 8hr. Then the cells were washed with PBS to remove contaminants. 2mL ethanol was added to the cells in order to open the membrane channel for the nucleus staining. The cells were rinsed again with PBS followed by addition of DAPI and 20 minutes incubation. The mixture was then washed 3 times with PBS. The image of the cells was obtained using confocal microscopy (CLSM, Leica, Germany).

RESULTS AND DISCUSSION

This synthesis of carbon dots from the crude dried latex of *C. gigantea* was conducted using pyrolysis. The process utilizes a muffle furnace for the probable dehydration and carbonization at a temperature as high as 300°C to aid the integration of its phytoconstituents [12]. At the initial stage after pyrolysis, the obtained dark brown powder was dissolved in sodium hydroxide solution to aid the completion of the CDs synthesis. After ultrasonication and filtration using 0.2µm filter membrane, a clear brown solution was obtained. The solution gave bluish light on exposure to UV lamp. The hydrodynamic diameter of the as-synthesized CDs was found to be around 8.7nm. Further characterizations that include FTIR and photoluminescence spectroscopy were conducted and the findings are reported in the next sections.

FTIR Analysis

FTIR analysis was recorded within the range of 400 to 4000 cm^{-1} (see Fig. 1). The FTIR spectrum exhibits peaks at 3250, 2968, and 1650 cm^{-1} , which can be attributed to O-H/N-H, C-H (sp^3) and C=O stretching, respectively [13]. The FTIR peak at 1432 could be assigned to C-N vibration [13] or carboxyl group [14,15] while the peak at 1148 could be assigned to C-O bond stretching [15]. It was assumed that crude latex comprises various organic compounds containing heteroatoms including nitrogen and oxygen. Therefore, it might not be surprising if peaks suggesting the presence of nitrogen functionalized groups are seen. The peaks at 862 and 686 cm^{-1} could be assigned to aromatic out-of-plane C-H bending [16]. The peak at 2470 cm^{-1} could be assigned to a possible C=N stretching [15]. Therefore, the FTIR spectra indicates the presence of carboxylic groups and possibly nitrogen-containing functional groups at the surface of the CDs. This observation could explain the hydrophilic nature of G-CDs.

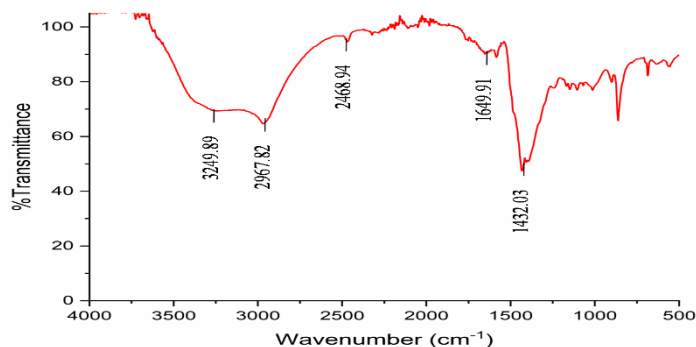


Figure 1. FTIR spectrum of G-CDs

Optical Property of G-CDs

In addition, the UV-Vis analysis of G-CDs showcased shoulder peak in the spectrum of G-CDs within 250-300 nm; this peak indicates $n-\pi^*$ transitions of G-CDs. Furthermore, the UV-Vis spectra showed band below 230 nm corresponding to $\pi-\pi^*$ transitions which occur within a pi-conjugated system (Fig. 2a).

Specifically, the absorption band below 230 nm is a result of the $\pi-\pi^*$ transition of C=C conjugates, while the absorption band at ~ 270 nm can be assigned to a $n-\pi^*$ transition of C=O and C=N vibrations on the surface of G-CDs [12].

PL analysis was carried out on G-CDs to characterize their photoluminescence property and quantum yield. The CDs showed PL spectrum when excited at an excitation wavelength of 320 nm. The PL showed broad emission whose maximum intensity was at a wavelength of 450 nm (Fig. 2b). The fluorescent QY obtained was about 21%. The broad emission could be as a result of randomized energy levels and complicated band structure [12]. The PL property of these CDs provides them with wide potential applications including cellular bioimaging.

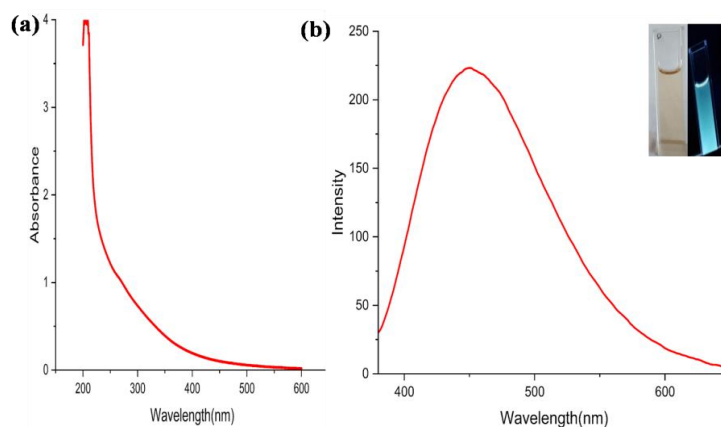


Figure 2. UV-vis spectra of G-CDs (a) and PL spectrum of G-CDs ($\lambda_{ex}=320$ nm). Inset: photographs of G-CDs under visible and UV light (b)

Colloidal stability (pH, ionic strength, and temperature effect)

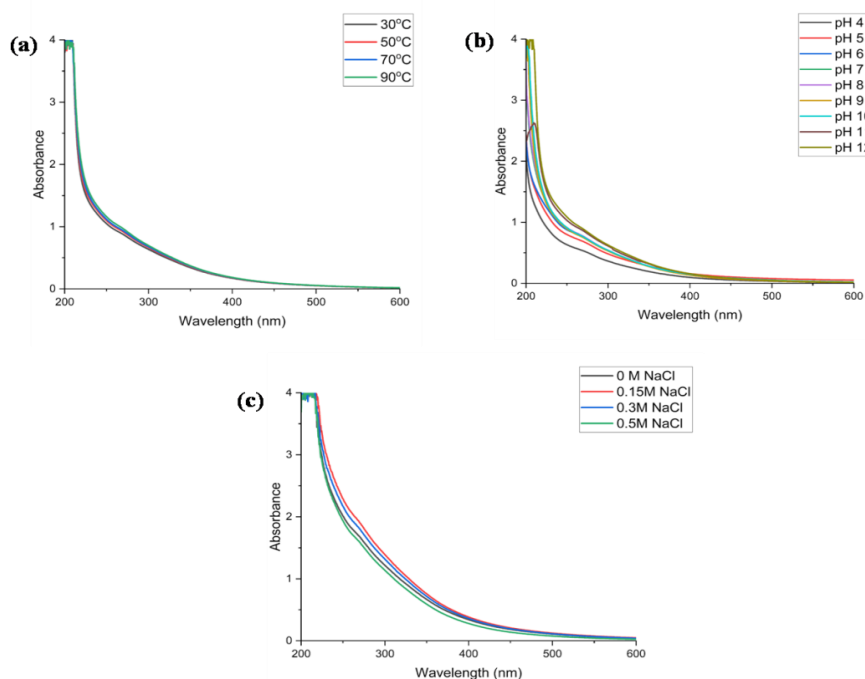


Figure 3. Colloidal stability at different conditions of (a) temperature (b) pH (c) ionic strength after 24 hrs.

Subsequently, stability of the G-CDs in colloids was investigated. Their stability at varied pH condition (4-12), varied temperature (30-90°C) and different ionic strength (0-0.5M NaCl) was studied. Influence of pH on the stability of G-CDs in solution was studied by adjusting the pH through addition of

0.1M NaOH and/or 0.1M HCl solutions to obtain the pHs of 4 up to 12. Subsequently, UV spectra of the varied solutions were taken at different time intervals of 1hr, 3 hr, 6hr and 24hr. As shown in figure 4b, the absorbance pattern of the G-CDs does not change significantly even at pH as low as 4 after 24hr. Similarly, after exposing the G-CDs to temperatures up to 90°C their UV-vis spectra remained comparably the same (Fig. 4a). Furthermore, a similar pattern was observed for the presumed G-CDs at different ionic strengths (Fig. 4c). This observation indicates the lack of change in the structural features of the presumed CDs thus indicating their ability to retain their property even at extreme conditions of temperature (as high as 90°C), pH (as low as 4) and ionic strength (up to 0.5M). Such property suggests the stability of presumed G-CDs at various conditions of the human body. Therefore, they could be suitable for biological applications [12].

Confocal Microscopy

In order to study the potency of new materials with fluorescence for cancer bioimaging, confocal microscopy is employed. Confocal microscope equipped with a laser was used to image breast cancer cells after being exposed with 200 ppm of our synthesized G-CDs for 8 hrs. Fluorescence of the presumed G-CDs served as an indicator for their potency in cellular imaging. Their colloidal stability indicated the possibility for biological applications. Here, we studied the ability of the CDs for bioimaging of MDA-MB-231 using confocal microscopy.

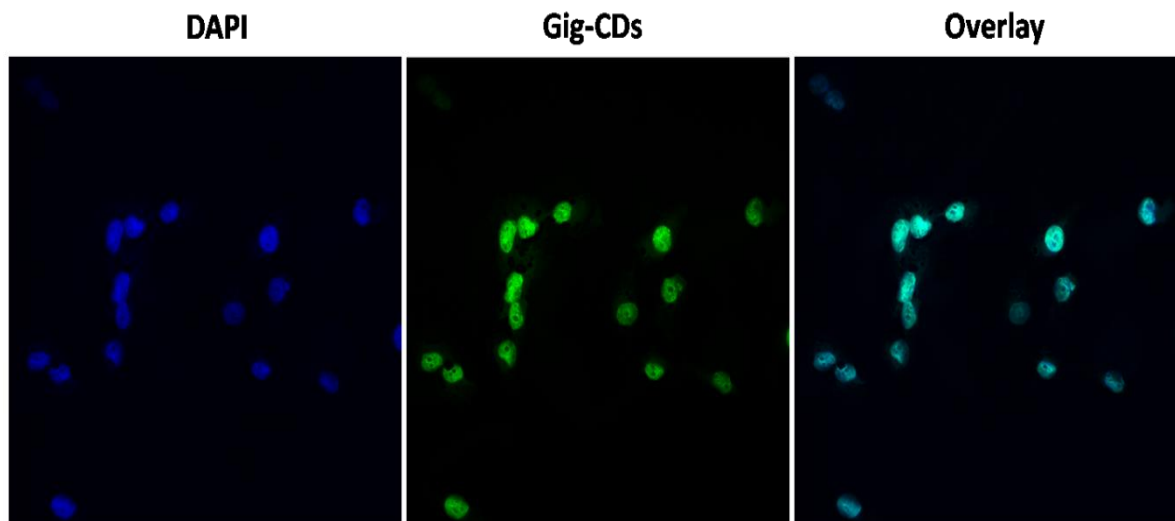


Figure 5. Confocal imaging of MDA-MB-231 breast cancer cell lines.

Figure 5 showed images taken of green fluorescence of G-CDs after laser irradiation after entering MDA-MB-231. Furthermore, the green fluorescence is seen in the cytoplasm of the cell whereas the nucleus (black dot surrounded by green) showed no obvious fluorescence. This indicates that the G-CDs are capable of occupying the cytoplasm. Catechin based CDs (cat-CDs) were used for imaging HeLa cell lines. Similar to our G-CDs, the Cat-CDs also gave green fluorescence [17]. In addition, Getachew *et al.* [18] used passivated metal halide perovskite quantum dots for imaging HeLa cells. Although metal based quantum dots were employed, the images obtained for the research were almost similar to the observation made in this study. As speculated somewhere else [19], the starting material of these CDs possessed targeted cytotoxic activity towards various types of cancer cells [8] thus it would not be surprising if the G-CDs took some of the targeting ability of its starting materials. Therefore, this calls for a detailed study into the biocompatibility and *in vivo* bioimaging application of these sustainable CDs.

CONCLUSION

In summary, we have synthesized fluorescent G-CDs from latex of *C. gigantea*. The CDs were characterized using UV-vis, FTIR, DLS, and PL. Furthermore, the CDs showed excellent colloidal stability at different conditions of pH, temperature and ionic strength. Finally, the CDs were able to enter into MDA-MB-231 breast cancer cells and exert fluorescence which is crucial for biological imaging.

ACKNOWLEDGMENT

The authors thank Universitas Airlangga for providing research facilities through the PFR research scheme under the contract number 1302/UN3.LPPM/PT.01.03/2023 and the Ministry of Education of the Republic of Indonesia.

REFERENCES

- [1] American Cancer Society, 80 (2022).
- [2] H. Puspitaningtyas, A. Espresso, S. H. Hutajulu, A. Fuad, and M. J. Allsop, *Cancer Control* **28**, 10732748211053464 (2021)
- [3] S. Masoudi Asil, E. D. Guerrero, G. Bugarini, J. Cayme, N. De Avila, J. Garcia, A. Hernandez, J. Mecado, Y. Madero, F. Moncayo, R. Olmos, D. Perches, J. Roman, D. Salcido-Padilla, E. Sanchez, C. Trejo, P. Trevino, M. Nurunnabi, and M. Narayan, *VIEW* **4**, 20220056 (2023).
- [4] H. Liu, Z. Li, Y. Sun, X. Geng, Y. Hu, H. Meng, J. Ge, and L. Qu, *Sci Rep* **8**, 1086 (2018).
- [5] N. Vasimalai, V. Vilas-Boas, J. Gallo, M. de F. Cerqueira, M. Menéndez-Miranda, J. M. Costa-Fernández, L. Diéguez, B. Espiña, and M. T. Fernández-Argüelles, *Beilstein J. Nanotechnol.* **9**, 530 (2018).
- [6] B. B. Karakoçak, A. Laradji, T. Primeau, M. Y. Berezin, S. Li, and N. Ravi, *ACS Appl Mater Interfaces* **13**, 277 (2021).
- [7] S. M. Bairagi, P. Ghule, and R. Gilhotra, 8 (2018).
- [8] S. Parhira, G.-Y. Zhu, M. Chen, L.-P. Bai, and Z.-H. Jiang, *Journal of Ethnopharmacology* **194**, 930 (2016).
- [9] B. Murugesan, J. Sonamuthu, N. Pandiyan, B. Pandi, S. Samayanan, and S. Mahalingam, *Journal of Photochemistry and Photobiology B: Biology* **178**, 371 (2018).
- [10] N. Sharma, I. Sharma, and M. K. Bera, *Journal of Fluorescence* **32**, 1 (2022).
- [11] K. J. Mintz, M. Bartoli, M. Rovere, Y. Zhou, S. D. Hettiarachchi, S. Paudyal, J. Chen, J. B. Domena, P. Y. Liyanage, R. Sampson, D. Khadka, R. R. Pandey, S. Huang, C. C. Chusuei, A. Tagliaferro, and R. M. Leblanc, *Carbon* **173**, 433 (2021).
- [12] Y. Y. Aung, A. N. Kristanti, S. Q. Khairunisa, N. Nasronudin, and M. Z. Fahmi, *ACS Biomater. Sci. Eng.* **6**, 4490 (2020).
- [13] Zhao, Wang, Zhao, Deng, and Xia, *Polymers* **11**, 1731 (2019).
- [14] S. Pandiyan, L. Arumugam, S. P. Srengan, R. Pitchan, P. Sevugan, K. Kannan, G. Pitchan, T. A. Hegde, and V. Gandhirajan, *ACS Omega* **5**, 30363 (2020).
- [15] K. Yadav, M. Das, N. Hassan, A. Mishra, J. Lahiri, A. K. Dubey, S. K. Yadav, and A. S. Parmar, *RSC Adv.* **11**, 9076 (2021).
- [16] A. F. Shaikh, M. S. Tamboli, R. H. Patil, A. Bhan, J. D. Ambekar, and B. B. Kale, *J Nanosci Nanotechnol* **19**, 2339 (2019).
- [17] Y. Kwee, A. N. Kristanti, N. S. Aminah, and M. Z. Fahmi, *Indones. J. Chem.* **20**, 1332 (2020).
- [18] G. Getachew, Y.-C. Tien, T.-C. Kan, W. B. Dirersa, A. Wibrianto, S. Orchirbat, J. Chang, A. S. Rasal, V. Gurav, S. Kizhepat, and J.-Y. Chang, *Chemical Engineering Journal* **467**, 143560 (2023).
- [19] M. A. Ahmad, Y. Aung, A. A. Widati, S. C. W. Sakti, S. Sumarsih, I. Irzaman, B. Yulianto, J. Chang, and M. Z. Fahmi, *Nanotheranostics* **7**, 187 (2023).

Nutrient Analysis of Compost from Vegetable and Fruit Waste and Its Application to Hydroponic *Ipomoea reptana* Plant

Refilda^{1*}, Andriani Febriyanti¹, Yefrida¹

¹Department of Chemistry, Faculty of Mathematics and Natural Sciences,
Andalas University, Padang, Indonesia.25161

*Corresponding author: refilda@sci.unand.ac.id

Abstract. Compost is the result of the decomposition of various organic materials by microbes which can be used as a source of nutrition for plants. Can it be used as a hydroponic medium? This study aimed to explore the potential of using vegetable and fruit waste as compost to be applied as hydroponic planting media for *Ipomoea reptana* plants. Compost was made for 30 days using the Takakura method with different compositions: 270 g of vegetable and fruit waste; 90 g of compost starter; EM-4 activator [0 mL (C1), 25 mL (C2), 50 mL (C3), 75 mL (C4), and 100 mL (C5)]; and 100 mL of water. The mixture was stirred once every two days. The physicochemical analysis of the compost revealed that the C5 composition met the compost standard according to SNI 19-7030-2004. The compost exhibited a blackish-brown color, crumbly texture, earthy smell, and contained approximately 1.15±0.10% nitrogen, 1.85±0.44% phosphorus, 0.93±0.05% potassium, 22.38±0.06% organic carbon, 0.09±0.01% iron, 0.02±0.04% zinc, with a pH of 7.46±0.03 and C/N 20.44±1.38%. The study also investigated the use of 1.5% compost as a hydroponic planting medium for *Ipomoea reptana* plants, employing the floating raft method. The results indicated that the C5 compost resulted in optimal growth for *Ipomoea reptana*, with performance comparable to plants grown on AB mix media (used as a positive control) and superior to those planted in aqueous media (used as a negative control). Consequently, it can be concluded that compost derived from vegetable and fruit waste presents a viable alternative to replacing AB mix fertilizer in hydroponic cultivation of *Ipomoea reptana* plants.

Keywords: Compost, EM-4, Hydroponics, *Ipomoea reptana*, Organic waste

INTRODUCTION

As the population continues to grow, the volume of waste increases, if left unchecked it will become a source of disease and very dangerous to health. Rapid population growth and urbanization rates have a disproportionate effect on the increase in waste [1]. In Indonesia, household organic waste consisting of waste vegetables, fruits, noodles, and others, constitutes the largest portion of waste, not only in 2022 but also in the previous year [2]. Composting offers a solution to minimize waste accumulation by converting household organic waste into organic fertilizer, as an alternative to the continuous use of synthetic fertilizers, which can negatively impact both soil and plant [3]. The nutrient content in synthetic fertilizers can be directly absorbed by plants to speed up the planting period. However, the use of synthetic fertilizers in the long term will harm soil productivity, leading to land degradation and reduced agricultural output [4].

Compost, derived from the decomposition of organic matter by microorganisms, serves as a valuable fertilizer. Studies by Kuchel *et al.* (2019) have demonstrated that composting effectively transforms organic waste into valuable products like organic fertilizer [5]. Furthermore, Cerda *et al.* (2018) have shown that biological fertilizers derived from household organic waste can be used as soil conditioners, reducing the reliance on chemical fertilizers, improving soil quality, and remediating contaminated soil, thus enhancing agricultural production [3]. The recommended composting method for individual or household use is the Takakura method, because this method can be applied on an individual or household scale in a simple and practical manner.

The composting process usually takes about 2-3 months requires the addition of bio decomposers, such as EM-4, which facilitate the breakdown of organic matter. EM-4 serves multiple functions, including improving soil structure and texture, supplying essential nutrients to plants, and inhibiting pest growth [4]. According to Gani *et al.* (2021), compost enriched with EM-4 demonstrates higher nutrient content in terms of nitrogen, phosphorus, and potassium compared to compost without the addition of EM-4 [5]. However, a previous study conducted by Della *et al.* (2020) on compost production from household organic waste using an EM-4 activator did not vary the addition of EM-4 and did not perform parameter tests such as determining N, P, and K [6]. Nurjazuli *et al.* (2019) found that composting is an effective approach to address waste-related issues. Composting has been widely practiced in the cultivation of hydroponic plants as well as traditional

farming, making it a convenient method to formulate and customize compost according to the specific requirements of plants [8].

Based on this information, the researchers are interested in developing compost using the Takakura method with varied EM-4 addition using household organic waste and applied as plant nutrition in hydroponic system. The resulting compost will be assessed for changes in color, texture, and odor. Chemical parameters such as C-organic content (determined using the Walkey and Black method), nitrogen (N, using the Kjeldahl method), phosphorus (P, using UV-Vis Spectrophotometry), potassium (K), iron (Fe), and zinc (Zn) (measured using Atomic Absorption Spectrophotometry, AAS) will be analyzed. pH levels will also be determined. The C/N value will be calculated and compared to the compost quality standards outlined in SNI. Furthermore, the produced compost will be applied to watercress plants hydroponically at a concentration of 1.5% using the floating raft method. The study will include 5 treatments and 3 replications, employing the Completely Randomized Design (CRD) experimental method. The effect of adding compost with varying concentrations of EM-4 on the growth of *Ipomoea reptana* plants will be observed in terms of plant height and leaf count. The data analysis will involve the one-way ANOVA test followed by the Duncan test.

The objective of this study is to determine optimal conditions for producing compost from household organic waste using the EM-4 activator. The research aims to assess the nutrient content of C-organic, N, P, K, Zn, Fe, C/N ratio, and pH, along with the physical properties of the resulting compost. Additionally, the study seeks to investigate the impact of varying concentrations of EM-4 in the compost-making process on the growth of *Ipomoea reptana* plants in a hydroponic system.

MATERIALS AND METHODS

Materials

The tools utilized in this study included knives, ovens, Takakura baskets, net pots, cloth covers, gloves, husk pads, used cardboard, spray flasks, filter paper (w-41), distillation units, analytical balances, Kjeldahl flasks, reagent bottles, desiccators, pH meters, styrofoam, UV-Vis Spectrophotometer, Atomic Absorption Spectrophotometer (SSA) (Varian Spectra AA-240), and other laboratory glassware.

The materials used in this study consisted of fruit and vegetable waste, Effective Microorganisms-4 (EM-4), distilled water, concentrated sulfuric acid (Merck), potassium dichromate (Merck), nitric acid (Merck), boric acid (Merck), sodium hydroxide (Merck), ammonium heptamolybdate tetrahydrate (Merck), potassium chloride (Merck), ascorbic acid (Merck), bromine cresol green, glucose (Merck), methyl red (Merck), potassium antimonytartarate (Merck), concentrated hydrochloric acid (Merck), peroxide acid (Merck), rockwool, AB-Mix nutrition, compost starter (Beraskin), and *Ipomoea reptana* seeds.

Producing Compost

The process of compost production involved the collection and cutting of fruit and vegetable waste into small pieces. Takakura baskets were prepared with a husk pillow placed at the bottom, and the inside of the basket was lined with cardboard. The basket was then filled with a mixture of finished starter/compost and household organic waste at a ratio of 1:3. EM-4 was added to the mixture. Husk pads were inserted, and the basket was covered with a cloth and a basket cover to prevent insects and flies from entering. The compost was stirred every 1 to 2 days. To monitor the composting process, hands were placed 2 cm away from the compost. If it felt warm, it indicated that the composting process was functioning properly. The composition of each compost is provided in Table 1.

Table 1. Composition of each compost

Sample Code	Fruit and vegetable waste (g)	Compost Starter (g)	EM-4 addition (mL)	Water (mL)
C1	270	90	0	100
C2	270	90	25	100
C3	270	90	50	100
C4	270	90	75	100
C5	270	90	100	100

Compost Nutrient Analysis

Destruction Stage

A sample of 1.0005 g of compost was placed in a Kjeldahl flask. Then, 10 mL of HNO₃ and 1 mL of H₂O₂ were added to the flask. The mixture was left for 30 minutes and then heated, starting at a temperature of 100°C. As the yellow steam dissipated, the temperature was increased to 200°C. The digestion process was considered complete when white steam and liquid were emitted from the clear flask. The flask was then allowed to cool, and the contents were diluted with distilled water and adjusted to a volume of 25 mL. The solution was stirred until homogeneous and filtered using W-41 filter paper to obtain a clear extract, which was used for the determination of P, K, Fe, and Zn.

Nitrogen (N)

0.5017 g of the sample was accurately weighed into a Kjeldahl flask. Then, 0.5 g of a selenium mixture and 10 mL of H₂SO₄ (pa) were added to the flask. The mixture was shaken vigorously until it became uniform. The complete digestion process was carried out gradually, starting at a temperature of 150°C and gradually increasing to a maximum temperature of 350°C. This resulted in obtaining a clear liquid. After cooling, the liquid was diluted with distilled water in a 50 mL flask, creating an extract. For the distillation process, 10 mL of the extract, 40 mL of distilled water, and 10 mL of 40% NaOH were pipetted into a Kjeldahl flask. A distillate container was prepared, consisting of 10 mL of 3% boric acid in a 100 mL Erlenmeyer flask, to which 3 drops of Conway indicator were added. Distillation was considered complete when the volume of the solution in the Erlenmeyer flask reached 60 mL. The distillate obtained from the process was then titrated with 0.05 N HCl until reaching the endpoint, indicated by a color change in the solution from green to light pink.

Phosphorus (P)

Calibration curves were prepared for each parameter using various concentrations. The concentrations used for each calibration curve were as follows:
Phosphorus (P): 0, 8, 16, 24, 32, 40, 48, 56 mg/L. Each standard series was pipetted with 1 mL, followed by the addition of 9 mL of phosphate dye reagent. The mixture was stirred until homogeneous, and the absorption was measured using a UV-Vis Spectrophotometer at a wavelength of 712 nm.

Potassium (K)

A calibration curve was created with concentrations of 0, 0.5, 1, 1.5, 2, and 2.5 mg/L. These concentrations were measured using an Atomic Absorption Spectrophotometer. To measure the sample solution, 1 mL of the extract was pipetted into a 10 mL volumetric flask. Then, 1 mL of the diluted extract was further pipetted into a 50 mL flask, and distilled water was added up to the mark. The solution was stirred until homogeneous and measured using an Atomic Absorption Spectrophotometer.

Fe

Calibration curves for Fe were constructed with concentrations of 0, 2, 4, 6, 8, and 10 mg/L. These concentrations were measured using an Atomic Absorption Spectrophotometer. For the sample solution, 1 mL of the extract was pipetted into a 10 mL volumetric flask and 9 mL of distilled water was added. The solution was stirred until homogeneous and measured using an Atomic Absorption Spectrophotometer.

Zn

The Zn calibration curve was prepared with concentrations of 0, 0.2, 0.4, 0.6, 0.8, and 1 mg/L, which were measured using an Atomic Absorption Spectrophotometer. For the sample solution, 1 mL of the extract was pipetted into a 50 mL volumetric flask. Distilled water was added up to the mark, and the solution was stirred until homogeneous. It was then measured using an Atomic Absorption Spectrophotometer.

C-Organic

The C-Organic calibration curve (Curve C) was generated with concentrations of 0, 100, 200, 300, 400, and 500 mg/L. For each sample, 1 mL of K₂Cr₂O₇ 2 N was added, followed by 1.4 mL of concentrated H₂SO₄. The mixture was stirred and allowed to stand for 30 minutes. The absorption was measured using a UV-Vis Spectrophotometer at a wavelength of 586 nm, resulting in a carbon standard calibration curve. In the treatment sample, 0.0256 g of the sample was weighed and placed into a 25 mL volumetric flask. Then, 2.5

mL of $K_2Cr_2O_7$ 2 N was added and stirred. Next, 3.5 mL of concentrated H_2SO_4 was added, followed by stirring and a 30-minute incubation period. The solution was diluted with distilled water, filtered, and measured using a UV-Vis Spectrophotometer at a wavelength of 586 nm.

Compost Application on Ipomeoa reptana Plants

Seeding Ipomeoa reptana Seeds and Transferring to Netpots

Rockwool blocks measuring 3 x 4 x 2.5 cm were prepared and sprayed with water. These blocks were arranged in a tray with holes. The seeds of *Ipomeoa reptana* were sown into the holes of the rockwool, with 6 seeds per rockwool block. The tray was then placed in a sunny location, and after 3 days, the seeds were ready to be transplanted.

Preparation of Planting Media

Styrofoam was used as the planting container. The surface of the styrofoam was perforated in circular shapes, ensuring that the hole size matched the diameter of the netpots to be used. The nutrient solution was prepared by mixing water with compost at a concentration of 1.5%. Each treatment was assigned a symbol: C1, C2, C3, C4, and C5. Additionally, a control variable, denoted as C(-), contained only water without any added nutrients. The control variable C(+) included the nutritional AB-Mix. The planting containers were placed in a sunny location and protected from rainfall. The nutrient water in the containers was regularly checked and replenished as necessary.

Observations on the Growth of Ipomeoa reptana Plants

The growth of *Ipomeoa reptana* plants was monitored by measuring the number of leaves and stem height once a week. After 28 days, the plants were harvested.

Data Analysis

The experimental design employed in this study was a completely randomized design (CRD) with 5 treatments and 3 replications. Each treatment unit consisted of one plant and was labeled with the respective symbols: C(-) for water, C1, C2, C3, C4, C5, and C(+) for AB-Mix. The collected data were subjected to variance analysis using a one-way ANOVA test, followed by the Duncan's test for further comparisons.

RESULTS AND DISCUSSION

Compost Observation Characteristics

In this study, compost was produced using fruit and vegetable waste. The maturity of the compost was assessed based on physical changes such as color, texture, and smell. Visual observations were conducted along with the use of simple equipment. The findings from these observations are presented in Table 2.

The color change in the compost is attributed to the activity of microorganisms involved in the decomposition process of organic matter. The addition of EM-4 during composting has a positive effect on color change, resulting in a blackish-brown appearance. The specific materials used in the composting process also influence this color change.

Table 2 presents the observed changes in compost texture after undergoing the composting process. In the treatment without the addition of EM-4 (C1), the compost exhibited a coarse texture, whereas treatments with the addition of EM-4 (C2, C3, C4, and C5) showed a crumbly texture. This indicates that the control group had incomplete texture changes due to the absence of EM-4, requiring a longer time for organic matter decomposition. Conversely, the addition of EM-4 in treatments (C2, C3, C4, and C5) resulted in the development of a crumbly texture, which signifies the activity of microorganisms involved in the decomposition process. These findings are consistent with Ekawandani's research (2018), which highlights texture changes as one of the characteristics of mature compost [9].

Table 2. Characteristics of compost with the addition of various concentrations of EM-4.

Compost	Characteristics		
	Color	Texture	Smell
SNI	Blackish-brown	Crumb	Earthy smell
C1	Brown	Rough	Old wood smell
C2	Blackish-brown	Crumb	Earthy smell
C3	Blackish-brown	Crumb	Earthy smell
C4	Blackish-brown	Crumb	Earthy smell
C5	Blackish-brown	Crumb	Earthy smell

Table 2 also demonstrates changes in odor after 30 days of composting. In treatment C1, all compost samples exhibited an odor resembling old wood, while treatments C2, C3, C4, and C5 showed an earthy smell, indicating maturity and proper microbial activity. This aligns with the perspective of Wiyono (2020), stating that mature compost emits a distinct earthy smell [10]. Additionally, according to Djuarnani and Setiawan (2019), mature compost lacks unpleasant odors. During decomposition, EM-4 facilitates the breakdown of nitrogen bonds, specifically converting ammonia into free nitrogen. Bacteria utilize free nitrogen as a constituent for their body proteins, leading to a reduction in ammonia odor [11].

Based on the provided information, compost samples C2, C3, C4, and C5, which involved varying concentrations of EM-4, meet the compost standards outlined in SNI. Conversely, treatment C1 without the addition of EM-4 does not meet the SNI standards.

Nutrient Content of Compost

According to Wahyono *et al.* (2020), the growth of microbial cells relies on nitrogen as an essential element for their formation and maintenance. A higher nitrogen content accelerates the decomposition of organic matter since microorganisms involved in compost decomposition require nitrogen for their development [12]. Cerda *et al.* (2018) also emphasize the significance of nitrogen (N) as a key nutrient in compost, playing a vital role in promoting plant growth and giving leaves a green color. Insufficient nitrogen in the soil can disrupt plant growth and development, leading to reduced yields due to inadequate chlorophyll formation, which is crucial for the photosynthetic process. On the other hand, excessive nitrogen can hinder flower growth and plant fertilizations [13]. As depicted in Figure 1, the highest average nitrogen content was observed in C5, which involved the addition of 100 mL of EM-4. The nitrogen content in C5 was determined to be $1.17 \pm 0.09\%$ with a pH value of 7.46 ± 0.03 . It is noteworthy that the compost derived from all sample treatments met the standard compost requirements outlined in SNI, which specify a minimum nitrogen content of 0.40%.

Based on Figure 1, the average phosphorus content ranges between 1.11% and 2.20%, with the highest phosphorus content observed in treatment C5. This finding is consistent with the nitrogen content in the compost. The elevated phosphorus levels can be attributed to the presence of phosphorus in the raw materials used for composting, as well as the abundant microbial population during the composting process. As microorganisms decompose, they release phosphorus, thereby increasing its concentration in the compost. Additionally, it is possible that the microorganisms present in EM-4 contribute to the enhancement and acceleration of the composting process. This aligns with the viewpoint of Kaswinarni and Nugraha (2020), who suggest that the weathering process of microorganisms can influence phosphorus levels [14]. Moreover, according to Putra and Nuraini (2019), high phosphorus levels in compost can be attributed to the presence of slime microorganisms, which enhance phosphorus nutrient content in the soil [15]. Referring to the data presented in Figure 2, it is evident that the phosphorus content in all variations meets the standard compost requirements outlined in SNI, which stipulate a minimum phosphorus content of 0.10%.

According to I Made *et al.* (2020) The presence of the nutrient potassium in compost is due to the fact that a lot of potassium comes from organic matter [16]. Harimbi *et al.* (2021) reported, potassium (K) in compost originates from the basic compost ingredients, particularly vegetable forage, which already contain K elements. However, these K elements are initially in the form of complex organic matter. Throughout the decomposition process, complex organic matter breaks down into simpler forms, releasing potassium that can be readily absorbed by plants. Potassium plays a crucial role in photosynthesis, protein formation, and cellulose production, which contribute to strengthening plant stems. The higher the potassium level in the compost, the better it is for promoting healthy plant stem growth [17]. Dwijosaputro (2018) also highlights the significance of potassium as an essential nutrient for plants. Insufficient potassium can inhibit the process of photosynthesis, leading to yellowing leaves and the development of spots [18]. Referring to Figure 1, it can

be observed that the potassium (K) content in all variations meets the standard compost requirements specified in SNI, which mandate a minimum potassium content of 0.2%.

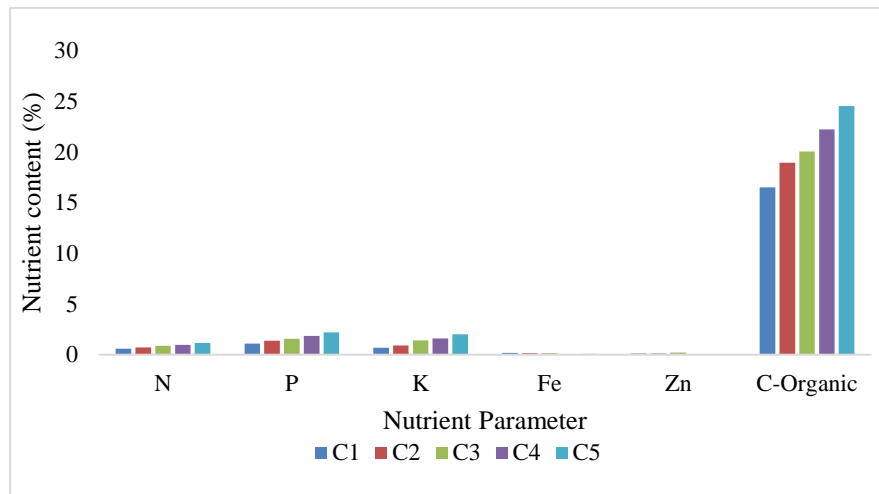


Figure 1. Graph Compost nutrient content with various variations of EM-4 addition.

The C-organic content in compost is influenced by the quality of organic matter and the activity of microorganisms involved in the decomposition process. In the analysis of organic carbon determination, the organic compound reduces orange Cr^{6+} to green Cr^{3+} in an acidic environment. As depicted in Figure 1, the highest average organic C content is observed in sample C5, which involved the addition of 100 mL of EM-4. This sample exhibited a C-organic content of $24.53 \pm 1.69\%$. The increased C-organic matter in compost can be attributed to the larger nutrient supply provided to microorganisms with the increasing amount of EM-4. Sujonto's research (2021) supports the significance of C-organic content in determining the mineral quality of compost, where higher total C-organic content signifies better mineral quality in the compost [19]. The C-organic content in all variations meets the standard compost requirements outlined in SNI, which specify a minimum range of 9.80% to 31%.

Iron (Fe) is an essential element that supports plant development and growth. It plays a crucial role in chlorophyll formation and photosynthesis. Figure 1 demonstrates that the Fe content in all variations meets the standard compost requirements outlined in SNI, which set a maximum Fe content of 2%. Zinc (Zn) is a vital element for plants, aiding in metabolic reactions, synthesizing plant growth compounds, and contributing to chlorophyll and carbohydrate production. According to Table 4.2, the Zn content in all variations meets the standard compost requirements specified by SNI, which establish a maximum Zn content of 0.5%.

Figure 2 illustrates the best growth of *Ipomoea reptana* plants, with the highest growth observed when provided with AB-mix nutrition followed by compost, and lastly, water alone. This demonstrates that hydroponically supplying compost to *Ipomoea reptana* plants yields better results compared to providing water alone. The favorable growth of water spinach plants given AB-mix nutrition can be attributed to its higher nutrient content compared to compost, leading to enhanced plant growth. This aligns with the findings of Sembiring's research (2023), which suggests that the application of AB-mix fertilizer outperforms N, P, K fertilizers and liquid compost fertilizers due to its complete array of elements that fulfill the plant's nutritional requirements [20]. AB-mix contains macro elements such as N, P, K, Ca, Mg, S, and microelements including Fe, Mn, Zn, and Cu.

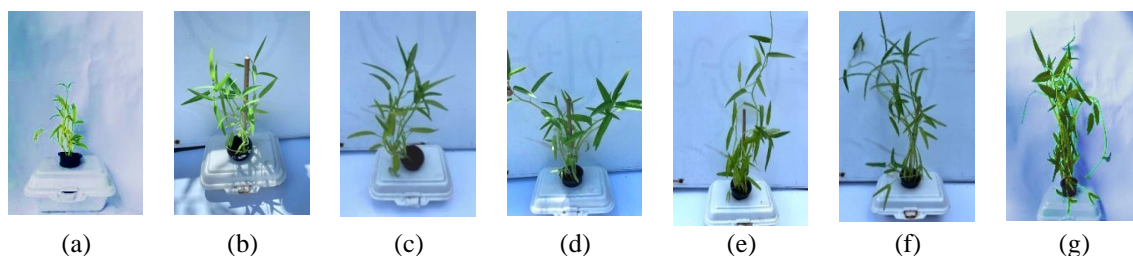


Figure 2. Comparison of the growth of *Ipomoea reptana* on growing media (a) Control (-) Water, Compost (b) C1, (c) C2, (d) C3, (e) C4 (f) C5, and (g) Control (+) AB-Mix.

The application of EM-4 in compost also influences the growth of *Ipomoea reptana* plants. As the amount of EM-4 added increases, the growth of *Ipomoea reptana* plants improves. This is evident from the obtained C/N ratio values, where a smaller C/N ratio indicates better growth of *Ipomoea reptana* plants. The highest-quality C/N ratio values were observed in C5 ($28.21 \pm 0.15\%$), C4 ($26.69 \pm 0.21\%$), C3 ($23.60 \pm 0.85\%$), C2 ($22.82 \pm 1.8\%$), and C1 ($20.86 \pm 0.04\%$).

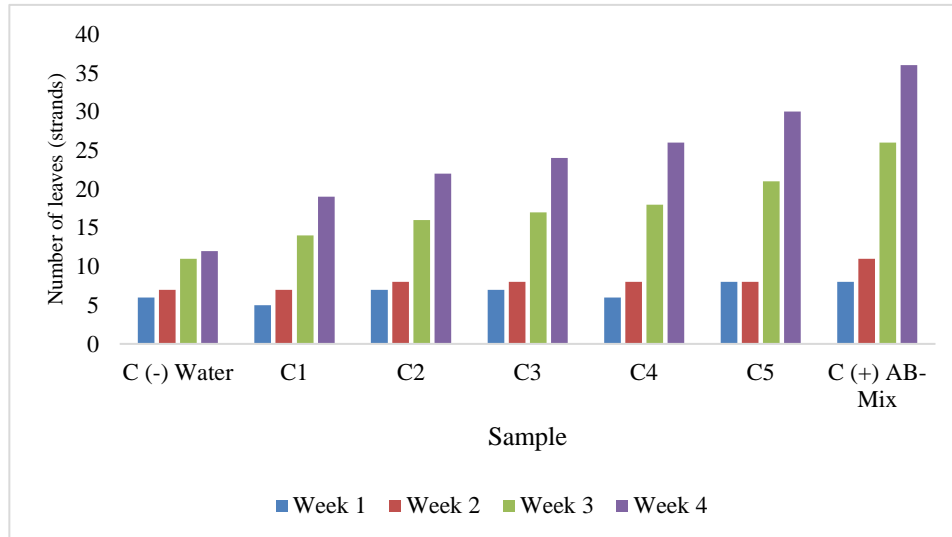


Figure 3. The average number of leaves of the *Ipomoea reptana* plant.

Based on Figures 3 and 4, it is evident that the number of leaves and height of *Ipomoea reptana* plants increase as the plants age. The highest average number of leaves and plant height are observed in week 5, particularly in the AB-Mix treatment, which served as the control group, with 36 leaves and a height of 35.67 cm. This is followed by the C5 treatment, which exhibited 30 leaves and a height of 28.33 cm. On the other hand, the lowest values were recorded in the Control (-) treatment, which had 12 leaves and a plant height of 12.67 cm. Notably, the application of compost with varying concentrations of EM-4 in the composting process significantly influenced the outcomes in each treatment, as observed in the leaf count at 7, 14, 21, and 28 days.

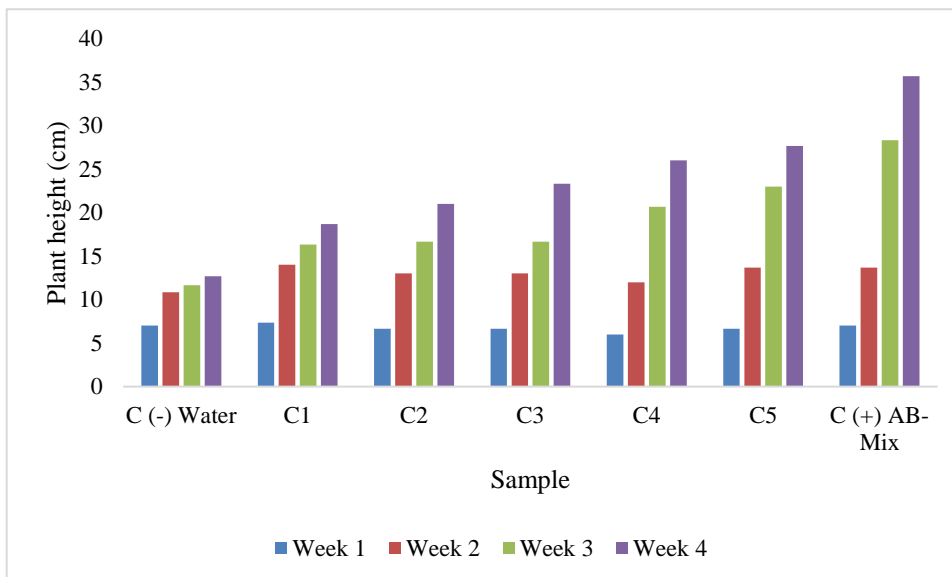


Figure 4. The average height of *Ipomoea reptana* plants.

CONCLUSION

The compost with the most favorable physicochemical properties among the compositions tested, which consisted of 270 g fruit and vegetable waste, 90 g starter, and varying amounts of EM-4 (0 mL for C1, 25 mL for C2, 50 mL for C3, 75 mL for C4, and 100 mL for C5), is compost C5. Compost C5 exhibits the following nutrient content: nitrogen $1.17 \pm 0.09\%$, phosphorus $2.20 \pm 0.08\%$, potassium $1.85 \pm 0.11\%$, C-organic $24.53 \pm 1.69\%$, Fe $0.09 \pm 0.01\%$, Zn $0.01 \pm 0.00\%$, pH 7.46 ± 0.03 , and C/N $20.86 \pm 0.04\%$. The application of 1.5% compost derived from various concentrations of EM-4 addition in composting significantly affects the growth of hydroponic *Ipomoea reptana*, particularly in terms of the number of leaves and stem height. Compost C5, which possesses the highest nutrient content, also yields the highest number of leaves and tallest stem height among the treatments.

REFERENCES

- [1] Nguyen, V. T. *et al.* Effects of C/N ratios and turning frequencies on the composting process of food waste and dry leaves. *Bioresour. Technol. Reports* **11**, 100527 (2020).
- [2] Cao, L. *et al.* Metagenomic analysis revealed the microbiota and metabolic function during co-composting of food waste and residual sludge for nitrogen and phosphorus transformation. *Sci. Total Environ.* **773**, 145561 (2021).
- [3] Sari, M. M., Ani, A. P. M. & Suryawan, I. W. K. Evaluation of the Implementation of Waste Generation Reduction Into Compost With Windrow System in the Talang Gulo Final Processing, Jambi City. *Konversi* **11**, 13–18 (2022).
- [4] Kumar, M., Ou, Y. L. & Lin, J. G. Co-composting of green waste and food waste at low C/N ratio. *Waste Manag.* **30**, 602–609 (2020).
- [5] Kucbel, M., Raclavská, H., Rzyikova, J., Vedova, B., Sassmanova, V., Drozdova, J., Raclavska, K., Juchelkova, D., 2019. Compost properties from home food waste generated ladder in an automatic composter. *J.Environment. Manage.* 236,657–666. <https://doi.org/10.1016/j.jenvman.2019.02.018>
- [6] Prinajati, P. D. Analysis of Household Waste Compost Fertilizer Processing as Plant Media. *J. ENVIROSAN* **1**, 68–71 (2018).
- [7] Cao, L. *et al.* Metagenomic analysis revealed the microbiota and metabolic function during co-composting of food waste and residual sludge for nitrogen and phosphorus transformation. *Sci. Total Environ.* **773**, 145561 (2021).
- [8] Song, C. *et al.* Impact of anti-acidification microbial consortium on carbohydrate metabolism of key microbes during food waste composting. *Bioresour. Technol.* **259**, 1–9 (2018)
- [9] Ekawandani, N. & Alvianingsih. Effectiveness of Leaf Compost. *Tedc* **12**, 145–149 (2018).
- [10] Wardani, K. The Quantity Effect of Earthworms (*Lumbricus Rubellus*) and Urea on The NPK Quality of Coffee Husk Waste Compost. **1**, 83–103 (2023).
- [11] I Made, Gede & Made (2020). Analysis of N, P, K Levels in Compost Fertilizer at Jagaraga Buleleng Landfill. Department of Chemistry: Faculty of Mathematics and Natural Sciences, Ganesha University of Education
- [12] Wardani, K. The Quantity Effect of Earthworms (*Lumbricus Rubellus*) and Urea on The NPK Quality of Coffee Husk Waste Compost. **1**, 83–103 (2023).
- [13] Cerda, A., Artola, A., Font, X., Barrena, R., Gea, T., Sanchez, A., 2018. Composting of food wastes: status and challenges. *Bioresour. Technol.* **248** (Pt A), 57–67
- [14] Giménez, A., Fernández, J. A., Pascual, J. A., Ros, M. & Egea-Gilabert, C. Application of directly brewed compost extract improves yield and quality in baby leaf lettuce grown hydroponically. *Agronomy* **10**, (2020).
- [15] Awasthi, M. K., Duan, Y., Awasthi, S. K., Liu, T. & Zhang, Z. Effect of biochar and bacterial inoculum additions on cow dung composting. *Bioresour. Technol.* **297**, 122407 (2020).
- [16] I Made, Gede & Made (2020). Analysis of N, P, K Levels in Compost Fertilizer at Jagaraga Buleleng Landfill. Department of Chemistry: Faculty of Mathematics and Natural Sciences, Ganesha University of Education
- [17] Korus, A. Effect of pre-treatment and drying methods on the content of minerals, B-group vitamins and tocopherols in *Ipomoea reptana* (*Brassica oleracea* L. var. *acephala*) leaves. *J. Food Sci. Technol.* **59**, 279–287 (2022).
- [18] Cerda, A., Artola, A., Font, X., Barrena, R., Gea, T., Sanchez, A., 2018. Composting of food wastes: status and challenges. *Bioresour. Technol.* **248** (PtA), 57–67

-
- [19] Lin, L., Xu, F., Ge, X., Li, Y., 2022. Improving the sustainability of organic waste management practices in the food-energy-water nexus: a comparative review of anaerobic digestion and composting. *Renew. Sust. Energ. Rev.* 89, 151–167.
- [20] Al Mussa Ugak, M. *et al.* Comparative study on passive aerated in-vessel composting of food wastes with the addition of Sabah ragi. *Carbon Resour. Convers.* 5, 200–210 (2022).

Effects of Kelakai (*Stenochlaena Palustris*) Extract on Exposure to Organophosphate Pesticide: In Silico and In Ovo Cytotoxic Studies

Nafisah¹, Hanna Habibah¹, Sarmila¹, Indah Saputri¹, Indah Setiawati¹ and Noer Komari^{1*}

¹Department of Chemistry, Faculty of Mathematics and Natural Sciences, Lambung Mangkurat University, Ahmad Yani Street, Banjarbaru City, 70714, Indonesia

*Corresponding author: nkomari@ulm.ac.id

Abstract. Vegetable farmers in South Kalimantan use pesticides to protect crops from pest attacks. The active ingredients of pesticides have widespread toxic effects on target and non-target organisms. Continuous exposure to pesticides causes cancer. Kelakai (*Stenochlaena palustris*) plant is suspected to have potential cytotoxic activity against cancer cell growth. This study aimed to examine the potential activity of chemical compounds in the protein Anaplastic Lymphoma Kinase (ALK) exposed to organophosphate pesticides in silico and in ovo. The in silico study was done using molecular docking and virtual screening methods. Kelakai methanol extract was obtained by maceration method. The in ovo study was carried out by injecting pesticide compounds and kelakai methanol extract at doses of 0.5 and 1 ppm to native chicken eggs. The results of in silico analysis showed that ethion and neophytadiene had the most negative ΔG values of -8.62 kcal/mol and -8.39 kcal/mol respectively, while the natural ligand 8 LY A 500 was -9.19 kcal/mol. The similarity in the type and number of residues in the binding complex between the ethion and neophytadiene ligands and ALK protein indicated the potential for competition between ethion and neophytadiene when binding to ALK protein. Neophytadiene is suspected as a potential anti-cancer compound by inhibiting the growth of ALK. Kelakai extract is considered capable of slowing down the rate of cell damage in chick embryos caused by ethion with its inhibitory ability so that the cell surface is not damaged quickly.

Keywords: Anaplastic Lymphoma Kinase, molecular docking, virtual screening, ethion, neophytadiene, in ovo.

INTRODUCTION

Vegetable farmers in South Kalimantan use pesticides to protect crops from pest attacks. Pesticides are used to eradicate pests and weeds or disturbing plants. In addition to their benefits, pesticides have the potential to poison and kill other living beings, including useful plants and insects, animals, and humans. The active ingredients in pesticides have widespread toxic effects, thus affecting target and non-target organisms, humans, as well as the environment and ecosystem as a whole [1]. Individual organisms exposed to pesticides have a greater chance of developing cancer. The most common cancers caused by pesticides are blood cancer, non-Hodgkins lymphoma, and brain cancer [2]. Cancer can be prevented by consuming vegetables [3].

Kelakai or kalakai (*Stenochlaena palustris*) is a fern type plant, which grows well in areas with high humidity such as peatlands in South Kalimantan [4]. Kelakai has a wide distribution area, has not been widely used and has no cultivation. This plant is used only for vegetables and is one of the proboscis monkeys' (*Larvatus nasalis*) food [5]. The content of kelakai compound is associated with many health benefits [6]. Kelakai leaf extract in the ethanol, n-hexane, and ethyl acetate fractions contains alkaloids, steroids, and flavonoids [7]. Extracts from the n-hexane fraction and ethyl acetate fraction of Kelakai herbs have cytotoxic activity against HepG2 liver cancer cell culture with an IC_{50} value of 865,84; 568,38; and 224,12 $\mu g mL^{-1}$ [8]. One of the extraction of molasses can be done by maceration (soaking) method. The maceration method is more efficient to obtain extract quantity, uses relatively little solvent, has a very minimal possibility of compound damage, and uses simple equipment [9].

Several studies have examined the relationship between pesticides and cancer risk [10,11], but the mechanism is not yet clear. The role of kelakai plants in cancer risk has also been studied [12,13], but still needs additional information. On the other hand, in ovo and in silico studies of the cytotoxicity of kelakai extract as an anticancer have never been studied. The in silico study aims to determine which organophosphate pesticide has the highest toxicity as a cause of cancer through virtual screening and to prove that the plant can be used as an anticancer agent using molecular docking. In ovo study was carried out on the growth of fertile native chicken eggs exposed to organophosphate pesticides and kelakai extract which aimed to examine the

effect of organophosphate pesticides with the highest toxicity on cell development in ovo and to prove that kelakai extract was able to inhibit the growth of cancer cells from organophosphate pesticides effectively in ovo.







MATERIALS AND METHODS

Materials

The laboratory equipment used included: rotary evaporator, water bath, syringe, surgical scissors, tweezers, beaker equipment, phone camera, and egg incubator. The hardware used was a laptop connected to the internet. The software used was Windows 11 64-bit, the USCF Chimera 1.16 program (<https://www.cgl.ucsf.edu/chimera/download.html>), Protein Data Bank database (<https://www.rcsb.org/>), PubChem database (<https://pubchem.ncbi.nlm.nih.gov/>), SwissDock web server (<http://www.swissdock.ch/>), pkCSM web server (<http://biosig.unimelb.edu.au/pkcsm/prediction>), and Prottox online tool web server (https://tox-new.charite.de/prottox_II/).

Materials for research included: samples of kelakai plant taken from Banjarbaru, South Kalimantan. Other ingredients are: n-hexane, methanol, 70% alcohol, ethion pesticide of the miton brand, distilled water, fertile native chicken eggs, 3D structure of the docking target protein, Anaplastic lymphoma kinase (ALK) The PDB 5USQ code and the original ligand 8LY A 500 were obtained from RCSB PDB. The structure of the organophosphate pesticide compounds and Kelakai compounds were obtained from Pubchem. Ten organophosphate compounds is shown in Table 1 and ten ligand compounds from Kelakai extract is presented in Table 2 [14].

Table 1. Compound of the active ingredients of organophosphate pesticides

No	Compound Name	Pubchem CID	Canonical Smiles	Structure
1	Diazinon	3017	<chem>CCOP(=S)(OCC)OC1=NC(=N C(=C1)C)C(C)C</chem>	
2	Parathion	991	<chem>CCOP(=S)(OCC)OC1=CC=C(C=C1)[N+](=O)[O-]</chem>	
3	Ethion	3286	<chem>CCOP(=S)(OCC)SCSP(=S)(OCC)OCC</chem>	
4	Profenofos	38779	<chem>CCCSP(=O)(OCC)OC1=C(C=C(C=C1)Br)Cl</chem>	
5	Malation	4004	<chem>CCOC(=O)CC(C(=O)OCC)SP(=S)(OC)OC</chem>	
6	Kloropirifos (Dursban)	2730	<chem>CCOP(=S)(OCC)OC1=NC(=C(C=C1)Cl)Cl)Cl</chem>	




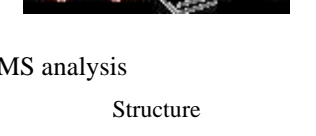
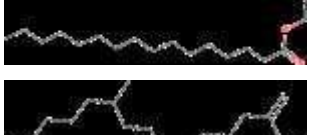
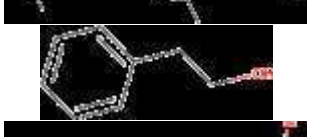
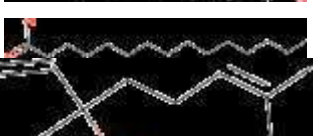


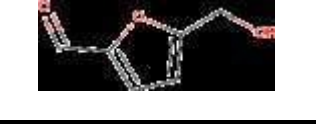
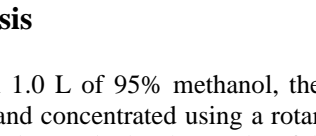


No	Compound Name	Pubchem CID	Canonical Smiles	Structure
7	Fenitrothion	31200	<chem>CC1=C(C=CC(=C1)OP(=S)(O)C)OC[N+](=O)[O-]</chem>	
8	Carbopheno-thion	13081	<chem>CCOP(=S)(OCC)SCSC1=CC=C(C=C1)Cl</chem>	
9	Dichlorvos	3039	<chem>COP(=O)(OC)OC=C(Cl)Cl</chem>	
10	Dicapthion (Dicaptan)	17168	<chem>COP(=S)(OC)OC1=C(C=C(C=C1)[N+](=O)[O-])Cl</chem>	

Table 2. Ligand compounds of kelakai methanol extract obtained from GC-MS analysis

No	Compound Name	Pubchem CID	Canonical Smiles	Structure
1	Hexadecanoic acid ethyl ester	12366	<chem>CCCCCCCCCCCCCCCC(=O)OCC</chem>	
2	Neofitadiena	10446	<chem>CC(C)CCCC(C)CCCC(C)CCC(C=C)C=C</chem>	
3	Phenethyl alcohol	6054	<chem>C1=CC=C(C=C1)CCO</chem>	
4	Hexadecanoic acid	985	<chem>CCCCCCCCCCCCCCCC(=O)O</chem>	
5	Pentadecanoic acid	13849	<chem>CCCCCCCCCCCCCCC(=O)O</chem>	
6	Linalool	6549	<chem>CC(=CCCC(C)(C=C)O)C</chem>	
7	Linoleic acid	5280450	<chem>CCC=CCC=CCC=CCCCCCC(=O)OCC(CO)O</chem>	
8	3-methyl acetate	6584	<chem>CC(=O)OC</chem>	
9	3-methyl-2,5-furandion (Citraconic anhydride)	12012	<chem>CC1=CC(=O)OC1=O</chem>	
10	5-hydroxymethyl-2-furancarboxal dehide	237332	<chem>C1=C(OC(=C1)C=O)CO</chem>	

Kelakai Extraction and GC-MS Analysis

Kelakai powder was weighed as much as 250.0 g and added with 1.0 L of 95% methanol, then macerated for 3 times 24 hours. The maserate was filtered to obtain filtrate and concentrated using a rotary evaporator to ¼ part of the initial solvent. Then, evaporation was done using and water bath. The results of the

methanol extract obtained were then analyzed to determine the type of compound information contained using GCMS QP 2010 (Shimadzu).

In Silico Analysis

ALK protein structure (PDB ID: 5USQ) was downloaded from the RCSB PDB site in *.pdb format. The protein was prepared using the UCSF Chimera 1.16 program by removing all residues. The natural ligand, 8LY was isolated from the protein and stored in *.mol2 format. The preparation of organophosphate pesticide ligands and the Kelakai compound ligands was also done using UCSF Chimera 1.16. Molecular docking was done using the SwissDock web server. The docking parameter taken was the Gibbs free energy value (ΔG). Visualization of the docking results was performed using the UCSF Chimera 1.16 program to show the type of ligand interaction and the residues of the target protein. Pharmacotoxic tests were carried out using the pkCSM and ProTox web servers.

In Ovo Activity Test

Fertile native chicken eggs as many as 24 eggs were divided into three groups as follows: group 1 without treatment (negative control), group 2 injected with organophosphate pesticides, and group 3 injected with organophosphate pesticides and kelakai methanol extract. Eggs were cleaned with running water and sterilized using 70% alcohol and put in an incubator. Embryo day 0 (E0) was the day when the eggs are placed in the incubator. Egg incubation was done at 37,8°C, humidity 55%. Group 1 embryo eggs were cracked at the embryo age of E0, E3, E7, E8, E9, E11, E14, E17, and E21. Eggs were transferred to a petri dish using tweezers and scissors to see the development of the embryo. Eggs from groups 2 and 3 were tested on the 9th day (E9) of incubation. The test solution was injected into the air cavity of the egg at a dose of 0.5 ppm and 1.0 ppm using a sterile syringe. After injection, the hole was sealed with plaster. Eggs were then observed after 6, 12, and 24 hours to see the symptoms of changes in blood cell growth and toxic symptoms in chicken embryos.

RESULTS AND DISCUSSION

Kelakai Extract and GC-MS Analysis

Kelakai (*Stenochlaena palustris*) was obtained by maceration method for 3 times for 24 hours. Kelakai methanol extract was obtained as much as 17.60 grams. Next, it was analyzed using GC-MS (Gas Chromatography Mass Spectrometry). The results of the analysis showed that there were 10 peaks from the methanol extract. The most amount of compound components in the methanol extract of Kelakai were located at peak 4 with a retention area value of 30.39% and a retention time of 32.026, with the possibility that the compounds present at peak 4 were 9-Octadecenoic acid (Z)- and methyl ester (CAS). The extraction results are shown in Figure 3. The methanol extract results were used for in ovo follow-up tests on exposure to fertile chicken eggs.



Figure 3. Kelakai methanol extract results

Molecular Docking

Target protein and ligand preparation for molecular docking was done using UCSF Chimera 1.16. Target protein was anaplastic lymphoma kinase (ALK) located on the short arm of chromosome 2 (2p23) and

the exchange of chromosomal segments with other genes. This protein contributes to the pathogenesis of various hematological malignancies and solid tumors, including lymphoma, lung cancer, inflammatory myofibroblastic tumor (IMT), Spitz tumor, renal carcinoma, thyroid cancer, gastrointestinal cancer, breast cancer, leukemia, and ovarian carcinoma [15]. The 3D ALK structure can be seen in Figure 1a. The 8LY A 500 ligand was taken from the target protein and prepared by minimizing energy. The 3D structure of the 8LY A 500 ligand can be seen in Figure 1b.



Figure 1. 3D structure: a. ALK protein, b. 8LY A 500 ligand

Molecular docking between ALK receptor proteins (PDB ID: 5USQ) (Figure 1a) with 10 pesticide compound ligands and 10 compound ligands of kelakai, as well as natural ligand 8LY A 500 (Figure 1b) was done using the SwissDock web server. The docking result parameter was the Gibbs free energy value (ΔG) and the interaction of the ligand bonds with amino acid residues. Table 3 shows the result of molecular docking between the ligand and the target protein. Gibbs free energy (ΔG) indicates how strong the bond is between the protein and the ligand, where the increasingly negative (lower) ΔG value indicates that the compound requires little energy when binding, so it can be interpreted that the compound has greater potential and activity to interact and form strong bonds with target proteins [8].

Table 3. Molecular docking results of ligands and ALK proteins

Ligands	ΔG (kkal/mol)	Amino Acid Residues	
		Hydrogen Bonding	Hydrophobic Interactions
Diazinon	-8.16	-	LEU260, GLY286, LEU340, ALA230, VAL219, SER280, ALA350, TYR249, PHE262, ASP351, LYS232, ILE211, ALA 230, GLU245.
Parathion	-7.68	-	TYR249, LEU260, ASP351, LEU340, GLY286, ASP281, TYR282, ALA230, SER280, VAL219, LYS232, LEU278
Ethion	-8.62	-	TYR282, GLY286, SER287, LEU340, ILE211, ALA230, VAL279, LYS232, LEU278, GLU245, ASP351, LEU260, SER280, VAL231, TYR282
Profenofos	-8.05	-	SER280, LEU340, LEU260, TYR282, ALA230, ALA350, VAL219, LYS232, ILE 211, ASP281, ASP351, GLY286, LYS337.
Malation	-7.95	-	SER280, LEU340, LEU260, TYR282, TYR249, PHE262, ALA230, VAL219, LYS232, VAL231, LEU278, GLU245, PHE262.
Kloropirifos (Dursban)	-7.63	-	GLY286, LEU340, TYR282, ALA350, ASP351, ILE211, VAL219, LYS232
Fenitrothion	-7.59	ASP351 (2.261 Å)	SER280, LEU340, TYR282, ALA230, ALA350, VAL219, LYS232, ILE 211, ASP281, ASP351, GLY286, LY232, TYR249.
Carbopheno-thion	-8.24	-	SER280, LEU278, TYR282, ALA230, VAL219, LYS232, ILE 211, PHE262, GLY286, TYR249, LEU260, LEU340.
Dichlorvos	-7.40	ALA380 (2.129 Å)	ALA403, ALA399, ASP400, LEU334, TYR378, THR375, ASP333, GLY374, VAL373, ARG332, VAL383, MET379
Dicaphthon	-8.16	ASP351 (2.426 Å)	ALA350, LEU340, GLY286, LEU260, ASP281, TYR282, ILE211, VAL219, ALA230, LYS232, TYR249.
Hexadecanoic acid ethyl ester	-7.08	LYS337 (1.878 Å)	ILE211, ALA230, ASP351, LEU340, ASN338.

Ligands	ΔG (kcal/mol)	Amino Acid Residues	
		Hydrogen Bonding	Hydrophobic Interactions
Neophytadiena	-8.39	-	ILE211, VAL219, LYS232, LEU278, ALA278, ALA230, TYR282, SER280, PHE262, TYR249, GLU284, LEU260, ASP351, GLY286, ALA350, LEU 340.
Phenethyl alcohol	-6.55	ARG332 (2.011 Å)	MET379, VAL383, THR375, ALA399, GLY374, VAL373, ASP400, LEU334, ALA403, TYR378.
Hexadecanoic acid	-8.09	-	ILE211, VAL219, LYS232, LEU278, ALA350, ALA230, TYR282, SER280, PHE262, TYR249, LEU360, ASP351, ASP290, GLY286, ALA350, LEU 340, ARG294.
Pentadecanoic acid	-8.00	-	ILE211, VAL219, LYS232, LEU278, ALA330, ALA230, TYR249, GLU245, LEU260, ASP351, SER280, GLY286, TYR282.
Linalool	-6.90	TYR378 (1.853 Å)	THR375, GLY374, ASP333, MET379, VAL383, ASP400, ALA403, ALA380, ALA399, ILE388, ARG372, VAL373.
Linoleic acid	-8.01	-	ILE211, VAL219, LYS232, LEU278, ALA230, TYR282, SER280, LEU260, PHE262, GLY286, LEU340, TYR249, SER287, ASP290, PHE289, LYS337.
3-methyl acetate	-5.88	-	MET379, VAL383, THR375, ALA399, ARG332, ASP333, LEU334, ALA403, TYR378.
3-methyl-2,5-furandion (Citraconic anhydride)	-6.54	THR375 (2.210 Å)	MET379, VAL383, THR375, ALA399, ARG332, ASP400, LEU334, ALA403, TYR378, GLY374, ASP333.
5-hydroxymethyl-2-furancarboxal dehide	-6.58	VAL373 (2.196Å) dan ALA403 (2.196 Å)	MET379, VAL383, THR378, ALA399, ARG332, ASP400, LEU334, TYR378, GLY374.
Native Ligand 8LY A 500	-9.19	-	ILE211, VAL219, VAL231, VAL279, LYS232, LEU278, ALA230, TYR282, SER280, LEU340, ASP281, PHE262, GLY286, ASP351, TYR249, GLU284, GLY286, ARG294.

Table 3 showed the results of molecular docking of 10 pesticide ligands, 10 Ligand compounds of kelakai methanol extract, and natural ligands of ALK protein. It was found that ethion and neophytadiene had the most negative ΔG values of -8.62 kcal/mol and -8.39 kcal/mol respectively. The more negative the ΔG value, the stronger the binding of the target protein-ligand complex due to the stability and strength of noncovalent interactions [16]. Ethion is thought to have more potential to damage ALK protein receptors than other pesticides. Although the ΔG value of neophytadiene is still higher than that of the natural ligand (8LY A 500) of the target protein, this ΔG value showed a value that is not too different which still allowed neophytadiene to act as a substitute for natural ligand and as a competitive inhibitor.

Neophytadiene was predicted to have anticancer activity against cancer growth through the inhibition of ALK protein. ALK played a role in the delivery process of signal transduction. This process began with the stimulation of kinases on the cell surface, followed by binding between kinases (dimerization). This dimerization process was followed by phosphorylation which is the process of transferring phosphate groups from one protein to the next in the signal transduction pathway. This transfer of phosphate group was the key to controlling the flow of signal delivery from the surface to the cell nucleus, which is the command for the cell to carry out the processes of proliferation (growth), division (cell division), differentiation (maturation), apoptosis (death), or metastasis (spread). ALK protein mutations caused autophosphorylation activity in signal transduction pathways and so cells continuously caused cancer growth. ALK mutations can occur in lung cancer, namely in neuroblastoma cancer cells, anaplastic large cell lymphoma, and NSCLC [11].

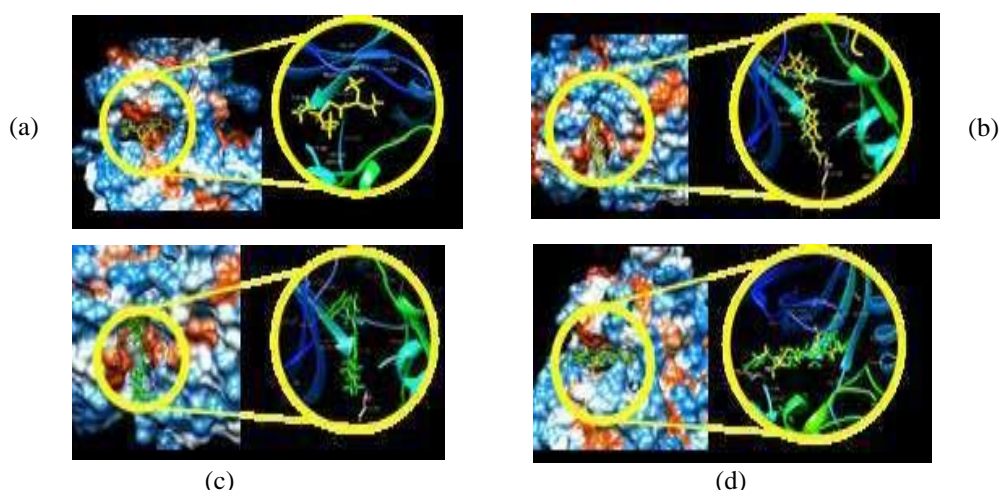


Figure 2. Molecular docking results with Anaplastic lymphoma kinase (ALK) protein (a) Ethion; (b) Neophytadiene (c) 8LY A 500; (d) Neophytadiene and ethion.

Figure 2 shows the visualization of molecular docking results. The ethion ligand was bonded to 15 residues of the ALK protein, which are: TYR282, GLY286, SER287, LEU340, ILE211, ALA230, VAL279, LYS232, LEU278, GLU245, ASP351, LEU260, SER280, VAL231, TYR282. Neophytadiene was bonded with 16 residues of ALK, which are ILE211, VAL219, LYS232, LEU278, ALA278, ALA230, TYR282, SER280, PHE262, TYR249, GLU284, LEU260, ASP351, GLY286, ALA350, LEU 340. The binding sites for ethion and neophytadiene compounds with ALK protein are right on the active site which had a similar interaction between the natural ligand 8LY A 500 amino acid residues and ALK protein. Binding site (active side) is an area on a protein that will be a place where ligands and proteins bind. The similarity of the type and number of residues in the binding complex between ethion and neophytadiene ligands and ALK protein indicated the potential for competition between ethion and neophytadiene when binding to ALK protein.

Virtual pharmacotoxic test was done using the pkCSM and Protox web servers with the parameters of Ames Toxicity, Hepatotoxicity, Skin Sensitization, LD₅₀Acute (mg/kg), and Class. The results of the analysis are shown in Table 4.

Table 4. Analysis results of pharmacokinetic prediction and ligand toxicity

Compound Name	Ames Toxicity*	Hepatotoxicity*	Skin Sensitisation*	LD ₅₀ Acute** (mg/kg)	Class**
Diazinon	No	Yes	No	17	2
Parathion	No	No	No	19	1
Ethion	No	No	No	13	2
Profenofos	No	No	No	162	3
Malation	No	No	No	190	3
Kloropirifos	No	No	No	60	3
Fenitrothion	Yes	No	No	229	3
Carbophenothion	No	No	No	7	2
Dichlorvos	Yes	No	Yes	17	2
Dicapthon	Yes	No	No	284	3
Hexadecanoic acid ethyl ester	No	No	Yes	5000	5
Neophytadiena	No	No	Yes	5050	6

Compound Name	Ames Toxicity*	Hepatotoxicity*	Skin Sensitisation*	LD ₅₀ Acute** (mg/kg)	Class**
Phenethyl alcohol	No	No	Yes	800	4
Hexadecanoic acid	No	No	Yes	900	4
Pentadecanoic acid	No	No	Yes	900	4
Linalool	No	No	Yes	2200	5
Linoleic acid	No	No	Yes	10000	6
3-methyl acetate	No	No	No	5000	5
3-methyl-2,5-furandion (Citraconic anhydride)	No	No	Yes	2600	5
5-hydroxymethyl-2- furancarboxal dehide	No	No	No	2500	5
Native ligand	No	Yes	No	200	3

Toxicity prediction aimed to predict the level of toxicity of compounds on the human body. Toxicity parameter LD₅₀ was classified based on Globally Harmonized System (GHS) divided into toxicity classes of I to VI. These six toxicity classes used LD₅₀ thresholds of 5, 50, 300, 2000, and 5000 mg/kg body weight with details as follows: class I (fatal if ingested) with a range of LD₅₀ ≤ 5 mg/kg, class II (fatal if ingested) with a range of 5 < LD₅₀ ≤ 50 mg/kg, class III (toxic if ingested) with a range of 50 < LD₅₀ ≤ 300 mg/kg, class IV (harmful if ingested) with a range of 300 < LD₅₀ ≤ 2000 mg/kg, class V (may be harmful if ingested) with a range of 2000 < LD₅₀ ≤ 5000 mg/kg, and class VI (not toxic) with a range of LD₅₀ > 5000 mg/kg. The Ames toxicity test was analyzed to see the mutagenic potential of the compound with bacteria. If a positive compound shows Ames toxicity, it is mutagenic and may act as a carcinogen [17].

Table 4 showed toxicity test using pkCSM Online Tools where ethion and neophytadiene were proven to have no toxicity in ames toxicity and hepatotoxicity. Neophytadiene in skin sensitization analysis showed positive results, so it was considered to cause skin sensitization. Natural ligand (8LY A 500) did not have toxicity in ames toxicity and skin sensitization analysis, but had hepatotoxicity so it can be predicted that the natural ligand was toxic to the target protein (ALK). Toxicity analysis of lethal dose 50 (LD₅₀) indicated that Ethion compounds had LD₅₀ value of 13 mg/kg, so it was classified in class 2 (fatal if ingested): 5 < LD₅₀ ≤ 50 mg/kg, which meant that the compound had a high toxicity effect. Therefore, Ethion has activator activity towards ALK protein. Natural ligand (8LY A 500) has a LD₅₀ value of 200 mg/kg, so it was classified in class 2 (fatal if ingested): 5 < LD₅₀ ≤ 50 mg/kg which meant that the compound had a high enough toxicity effect. Meanwhile, the neophytadiene compound had a LD₅₀ value of 5050 mg/kg, so the compound was classified in class 6 (not toxic): LD₅₀ > 5000 mg/kg, which meant that it had a really low level of toxicity. The value of LD₅₀ shows that the greater the toxic value, the less toxic a compound is, vice versa [18].

In Ovo Cytotoxic Activity Test

Figure 4 shows the development of native chicken embryos from 0 days to 21 days. Figure (4a) shows the embryo on day 0 where in the white center is a blastoderm zygote cell, then the blastoderm has red stripes which are an indication of the start of the blood circulation system (Figure 4b). The development of the embryo can be seen from the elongated shape of the center of the blastoderm [19]. Incubation of more than 7 days showed that limbs have been formed, eyes have protruded, and limbs have been formed (Figure 4c). Embryonic organ growth was complete and increasingly showed the development of the organ structure of the chicken embryo. This is reinforced by the statement of Kusumawati et al. (2016) that the organs of chicken embryos in incubation of more than 7 days are sufficiently complete. Incubation period of more than 10 days shows visible brain, beak, feathers, and wings (Figures 4d, 4e, and 4f). Growth towards organ perfection occurred on day 15 and beyond (Figures 4g, 4h, and 4i). Negative control observations (without treatment) showed good results from 0 days to 21 days until chicks ready to hatch were formed. Embryonic development proceeded normally and there were no growth disturbances.



Figure 4. The results of negative control test (a) embryo 0 hari (b) embryo 3 hari (c) embryo 8 hari; (d) embryo 9 hari; (e) embryo 10 hari; (f) embryo 12 hari; (g) embryo 15 hari; (h) embryo 18 hari; (i) embryo 21 hari.



Figure 5. The results of ethion activity test (E) dan ekstrak kelakai methanol (M) (a) E 0,5 ppm, 6 jam; (b) E 0,5 ppm, 12 jam; (c) E 0,5 ppm, 24 jam; (d) E 1 ppm, 6 jam; (e) E 1 ppm, 12 jam; (f) E 1 ppm, 24 jam; (g) ME 0,5 ppm, 6 jam; (h) ME 0,5 ppm, 12 jam; (i) ME 0,5 ppm, 24 jam; (j) ME 1 ppm, 6 jam; (k) ME 1 ppm, 12 jam; (l) ME 1 ppm, 24 jam.

Figure 5 is the observation of the development of chicken eggs injected with ethion 0.5 and 1 ppm. Figures 5a, 5b and 5c shows changes in the surface of embryo cells with lots of white spots, and so cell aggregation became irregular which resulted in abnormal development and even death of the embryo. In addition, exposure to ethion formed a residue in the yolk which will interfere with the growth and development of the chick embryo which may result in developmental abnormalities because the yolk is needed as a source of nutrition (Figures 5d, 5e and 5f). Thus, exposure to ethion showed a change in the embryo when compared to the development of the embryo without any injection. Furthermore, exposure to kelakai methanol extract injected together with ethion showed better results than ethion alone (Figures 5g, 5h and 5i). The addition of kelakai extract is thought to be able to compensate for cell damage caused by ethion with its inhibitory ability so that the cell surface is not too damaged, although there was still a small amount of ethion residue on the cell surface (Figures 5j, 5k, and 5l). This may also be strengthened that the neophytadiene and ethion compounds which were studied *in silico* using ALK target proteins showed that these two compounds binded 15-16 amino acid residues at the same binding site. So there was competition between the two to be on the active side of the target protein. Therefore, it is possible that ethion and neophyadiene have the same ability to bind to cells.

CONCLUSION

Ethion is the most toxic and potentially cytogenic organophosphate pesticide with $\Delta G = -8.62$ kcal/mol. Neophytadiene from kelakai has potential as a competitor in inhibiting Anaplastic Lymphoma

Kinase (ALK) protein with $\Delta G = -8.39$ kcal/mol. The similarity in the type and number of bond residues in ethion and neophytadiene in ALK protein indicates competition between ethion and neophytadiene. Exposure to ethion interferes with the growth and development of chicken egg embryos which results in embryonic death. Kelakai extract is considered capable of slowing down the rate of cell damage in chick embryos caused by ethion.

ACKNOWLEDGEMENT

Gratitude are conveyed to the Directorate General of Learning and Student Affairs, Ministry of Research, Technology, and Higher Education of the Republic of Indonesia via Program Kreativitas Mahasiswa (PKM) year 2022 as well as Universitas Lambung Mangkurat which have provided financial support for this research.

REFERENCE

- [1] G. V. Souisa, B. Talarima, and Rehena Z, Peningkatan Perilaku Pencegahan Dampak Pestisida Pada Kesehatan Petani. *J Pengabd Kpd Masy* **26**, 109-113 (2020).
- [2] O. S. Pamungkas, Bahaya Paparan Pestisida terhadap Kesehatan Manusia. *Bioedukasi* **14**, 27-31 (2017).
- [3] D. Marisa, L. Hayatie, S. Juliati, E. Suhartono, and N. Komari, Molecular docking of phytosterol compounds from kelakai (*Stenochlaena palustris*) as anti-breast cancer. *Acta Biochim Indonesia* **4**, 59-66 (2021).
- [4] M. H. Nurinayah, M. A. Soendjoto, and Darmono, "Jenis tumbuhan paku di Kawasan Rawa Sungai Lumbah, Kabupaten Barito Kuala" in *Prosiding Seminar Nasional Lahan Basah*, (2016) pp. 141-145.
- [5] S. Iskandar, H. S. Alikodra, M. Bismark, and A. P. Kartono, Status Populasi Dan Konservasi Bekantan (*Nasalis Larvatus* Wurm. 1787) di Habitat Rawa Gelam, Kalimantan Selatan. *J Penelit Hutan dan Konserv Alam* **14**, 123-132 (2017).
- [6] D. M. Maharani, S. N. Haidah, and Haiyinah, Studi Potensi Kelakai (*Stenochlaena palustris* (Burm.F) Bedd) Sebagai Pangan Fungsional. *PMKP-1-13-1* (2000).
- [7] Erwin, D. Anggeraini, and Suryani. Chemical analysis and antibacterial activity of the ethanolic extract of *Stenochlaena palustris*. *Der Pharm Lettre* **8**, 1-7 (2016).
- [8] M. N. Yanti, I. Rahmawati, and W. Herdwiani, Uji Aktivitas Sitotoksik Herba Kelakai (*Stenochlaena Palustris* (Burm.F.) Bedd.) Terhadap Sel Kanker Hati HEPG2. *Bioteknologi dan Biosanis Indonesia* **8**, 255-266 (2021).
- [9] A. Damayanti, and A. Fitriana, Pemungutan Minyak Atsiri Mawar (Rose Oil) Dengan Metode Maserasi. *J Bahan Alam Terbarukan* **1**, 1-8 (2012).
- [10] R. Agustin, and Muhartono. Dampak Penggunaan Pestisida Organoklorin terhadap Risiko Kanker Payudara. *J Agromedicine* **5**, 433-437 (2018).
- [11] N. N. Soeroso, and F. R. Ananda, Lung Cancer among Never-Smoker Women: An Epidemiological Data in North Sumatera, Indonesia. *Int J Respir Med* **1**, 1-9 (2019).
- [12] H. M. Mashar, and I. Annah, Cytotoxicity of Kelakai (*Stenochlaena palustris*) Extract to MCF-7 Breast Cancer Cell. *J Fitofarmaka Indonesia* **7**, 5-9 (2020).
- [13] D. Margono, E. Suhartono, and H. Arwati, Pengaruh Ekstrak Kelakai (*Stenochlaena Palustris* (Burm.F) Bedd) Terhadap Kadar Interleukin-10 (IL-10) Mencit. *Med Lab Technol J* **2**, 31-36 (2016).
- [14] N. Nurmilatina. Analisis Komposisi Kimia Daun Kelakai (*Stenochlaena palustris* Bedd.) dengan Berbagai Pelarut menggunakan GCMS (Chemical Composition Analysis of *Stenochlaena palustris* Bedd. Leaves using Various Solvents on GCMS). *J Ris Ind Has Hutan* **9**, 9-16 (2017).
- [15] Z. Cao, Q. Gao, Fu M, Ni NAN, Pei Y, and Ou W Bin, Anaplastic lymphoma kinase fusions: Roles in cancer and therapeutic perspectives (review). *Oncology Letters* **17**, 2020-2030 (2019).
- [16] A. Suhadi, R. Rizarullah, and F. Feriyani, Simulasi Docking Senyawa Aktif Daun Binahong Sebagai Inhibitor Enzyme Aldose Reductase. *Sel J Penelit Kesehatan* **6**, 55-65 (2019).
- [17] F. Naufa, R. Mutiah, Y. Yen, and A. Indrawijaya, Studi in Silico Potensi Senyawa Katekin Teh Hijau (*Camellia sinensis*) sebagai Antivirus SARS CoV-2 terhadap Spike Glycoprotein (6LZG) dan Main Protease (5R7Y). *Journal of Food and Pharmaceutical Sciences* **10**, 584-596 (2022).
- [18] A. F. Faqiha, Y.Y.A. Indrawijaya, A. Suryadinata, Amiruddin M, and Mutiah R. Potensi Senyawa

Nitazoxanide dan Arbidol sebagai Antivirus SARS-CoV-2 terhadap Reseptor NSP5 (7BQY dan 2GZ7) dan ACE2 (3D0G dan 1R4L). *Journal of Food and Pharmaceutical Sciences* **10**, 570-583 (2022).

- [19] F. Fitriani, H. Husmimi, D. Masyitha, and M. Akmal, Histologis Perkembangan Embrio Ayam pada Masa Inkubasi Satu sampai Tujuh Hari. *J Agripet* **21**, 65-70 (2021).
- [20] A. Kusumawati, R. Febriany, S. Hananti, D. M. Sartika Dewi, and N. Istiyawati, Reproduksi dan Kebidanan Fakultas Kedokteran Hewan D, et al. Perkembangan Embrio dan Penentuan Jenis Kelamin DOC (Day-Old Chicken) Ayam Jawa Super. *J Sain Vet* **34**, 29-41 (2016).

The Rapid Identification Metabolites of *Aaptos* sp. from Bunaken National Park, Manado

Annisa Elcentia Fajarwati^{1,2}, Mutia Hardhiyuna¹, Febby Nurdiya Ningsih¹, Andi Setiawan², Ucu Yanu Arbi³, Dwi Wahyu Indriani¹, and Peni Ahmadi^{1,*}

¹ Research Center for Vaccines and Drugs, Life Sciences Research Organization, National Research and Innovation Agency (BRIN), Jl. Raya Bogor Km. 46, Cibinong, 16911, Indonesia

² Department of Chemistry, Faculty of Mathematics and Natural Sciences, Lampung University, Bandar Lampung 35145, Indonesia

³ Research Center for Oceanography, National Research and Innovation (BRIN), Jl. Pasir Putih I, Ancol Timur, Jakarta Utara, 14430, Indonesia

*Corresponding author: peni.ahmadi@brin.go.id

Abstract. The sponge *Aaptos* sp. has been found to contain a wide range of bioactive compounds. In this study, we used liquid chromatography-mass spectrometry (LC-MS/MS) to quickly identify bioactive metabolites in the water fraction of the sponge *Aaptos* sp. collected from Bunaken National Park, Indonesia. In this work, LCMS/MS coupled with quadrupole-time-of-flight mass spectrometry and MassLynx software was established to characterize the chemical compositions of the compounds. Based on the LC-MS/MS and ¹H-NMR analysis, aaptamine is the major compound detected in the water fraction of *Aaptos* sp. The anticancer bioactivity test was conducted using the microtetrazolium (MTT) method against colorectal cancer cells (Caco-2 and DLD-1) as well as normal cells (HEK293), while the antibacterial test used the TLC bioautography method against resistant clinical pathogenic bacteria *Staphylococcus aureus*. The results showed that water fraction has moderate cytotoxic activity as an anticancer agent against colorectal cancer cells Caco-2 and DLD-1, but did not display cytotoxic activity against normal cells HEK293. Additionally, the fraction demonstrated antibacterial potential against resistant clinical pathogenic bacteria *S. aureus*. These findings confirmed that water fraction from *Aaptos* sp. contained aaptamine and can be used as an antibacterial inhibitor as well as an anticancer agent. This research also demonstrated the importance of rapidly identifying bioactive compounds using LCMS/MS supported by MassLynx database to accelerate the discovery of compounds and holds promise strategy for future drug development.

Keywords: *Aaptos* sp., aaptamine, clinical-resistant *Staphylococcus aureus*, anticancer, bioautography

INTRODUCTION

Sponges (Porifera) are known to contain a wide variety of secondary metabolites which are promising pharmaceutical agents [1]. The genus *Aaptos* is one of the most widespread sponge species in the world's marine ecosystems, including in Indonesia. The investigation of genus *Aaptos* showed that aaptamine alkaloids are a group frequently encountered. Aaptamine has a unique 1H-benzo[de]-1,6-naphthyridine skeleton [2]. Various aaptamine have been isolated from *Aaptos* such as aaptamine from *Aaptos suberitoides* [3], 3,5-dicarbomethoxy-1,6-naphthyridine from *Aaptos aaptos* [4], and 2-methoxy-3-oxoaaptamine from *Aaptos* sp. [5]. Therefore, these metabolites are the great interest to scientists due to their biological properties such as cytotoxic, antiviral, antibacterial, antifungal, and antioxidant [6].

To date, many research groups are conducting chemical analyses on *Aaptos* in different regions to investigate their structural diversity and bioactivity [7]. The investigation has been severely hindered by a major challenge, namely their limited abundance. In marine invertebrates, many compounds are only present in concentrations lower than 0.00001% of their body weight [8]. The development of new techniques, such as LC-MS/MS offers the potential for identification of target components without the need for extensive purification and using only a small sample. LC/MS increases sensitivity and speeds up analysis. These spectra have wide applications in structure identification and elucidation. The improvements in LC/MS coupled with hybrid quadrupole/time-of-flight (TOF) geometries enable identification with a high degree of confidence [9]. In a similar attempt, LC-MS can be used as a tool for rapid dereplication for drug discovery featuring information on molecular weight with existing natural product databases [10].

In this study, we describe the chemical structure of aaptamine isolated from *Aaptos* sp collected from Bunaken National Park, Indonesia, using LC-MS/MS and combined with the data from ¹H-NMR. Their bioactivities as antibacterial against clinical-resistant *Staphylococcus aureus* and the cytotoxicity properties against colorectal cancer cells (Caco-2 and DLD-1), as well as normal cells (HEK293), are also reported. This

research also demonstrated the importance of rapidly identifying bioactive compounds using LCMS/MS supported by the MassLynx database to accelerate the discovery of compounds and holds promise strategy for future drug development.

MATERIALS AND METHODS

Animal Material

The sponge *Aaptos* sp. studied was obtained by hand at a depth of 10-15 meters from Likuan 3, Bunaken National Park, North Sulawesi, Indonesia in September 2022, and kept frozen until extraction. The specimen was deposited at Genomics Laboratory, National Research and Innovation Agency (BRIN), Cibinong, Bogor prior to extraction.

Extraction

The extraction method refers to Ahmadi *et al* [11]. The specimen sponge (80.94 g, wet) was thawed, chopped, and macerated using methanol (MeOH) until submerged overnight before being filtered. The procedure was repeated three times. All the filtrates were combined and concentrated using a rotary evaporator until thick extracts were obtained. Subsequently, the extract was partitioned exhaustively between H₂O and EtOAc. The water extract was dried using freeze-drying and the residue was washed with MeOH and evaporated under vacuum to give a soluble extract (5504.68 mg).

Thin Layer Chromatography (TLC)

TLC was performed using TLC sheets of silica GF₂₅₄ as stationary phase and dichloromethane (DCM)-methanol (MeOH) (9:1) as mobile phase. Silica plates were visualized under UV₂₅₄ to reveal a conjugated bonds compound. Further, the plates were reacted using several specific reagents to reveal spots of different groups using cerium sulfate (Ce (SO₄)₂), ninhydrin, and Dragendorff. Cerium sulfate reagent was used to determine the content in the sample by marking the appearance of blackish-brown stains. Dragendorff reagent was characterized by the formation of orange stains. Ninhydrin reagent was characterized by the formation of purple spots. The heated TLC plate was observed and the R_f value was calculated to determine the components of the compounds present in the extract.

Liquid Chromatography-Mass Spectrometry (LC-MS/MS)

The water extract was analyzed using Waters ACQUITY UPLC® H-Class System (Waters, Beverly, MA, USA), ACQUITY UPLC® HSS C18 column (11.8 μm, 2.1 Å, ~100 mm) (Waters, Beverly, MA, USA), and coupled with Xevo G2-S QTOF Mass Spectro (Waters, Beverly, MA, USA) with system electron spray ionization (ESI). The mass spectrometer was operated in a positive ion mode system. The resulting data is in the form of chromatogram (LC) and spectra (MS). The elemental compositions were calculated using Waters MassLynx software (v4.1). Further, the type of compound compared to database PubChem.

Chromatography Column (CC)

A portion of the water fraction (250 mg) was further fractionated using a chromatography column to reduce impurities in the sample. The extracts were fractionated using solvents of different polarities (100% aquadest, 25% MeOH, 50% MeOH, 75% MeOH, and 100% MeOH) on reversed-phase (RP-C18) column to result in the fractions of F1 (139.91 mg), F2 (31.05 mg), F3 (0.55 mg), F4 (1.77 mg), and F5 (7.35 mg).

Nuclear Magnetic Resonance (1H-NMR)

The major fraction of water extract (F1) was dissolved in protonated solvent MeOH-d₄ and analyzed on a JEOL spectrometer JNM-ECZ500R/S1 either at 500 MHz for ¹H-NMR.

Test Organism

The clinical pathogenic bacteria-resistant *Staphylococcus aureus* was obtained from deposited UPT-LTSIT, Lampung University, Bandar Lampung, Indonesia. The previous study showed that *Staphylococcus aureus* was resistant to amoxicillin (25 µg), ciprofloxacin (5 µg), and clindamycin (2 µg) [12].

Maintenance of Cultures of Pathogenic Bacteria

The pathogenic isolate of *Staphylococcus aureus* resistant kept in an incubator at 37 °C on Mueller Hinton Agar (MHA) slants was used for a month. Prior to each assay, the culture was revived in Mueller Hinton Broth (MHB) liquid medium, which was incubated at 37 °C in the incubator for 18 hours. Bacterial suspensions from tubes were then subcultured on MHB.

Bioautography Assay Using Staphylococcus aureus resistant

This study used the bioautographic TLC technique to visually determine the clear zone of bacterial growth inhibition. The bioautographic assay was carried out on TLC plates refers to Chomnawang et al. [13] with some modifications including mobile phase, reagent, and test bacteria. The water extract with a concentration of 0.5 mg/mL was subjected to thin-layer chromatography (TLC). Silica gel GF₂₅₄ plates were developed with n-Hexane: EtOAc (7:3). The plates were either visualized using Dragendorff. This study used the bioautographic technique to reveal the visually observing clear zones of bacterial growth inhibition for guiding the active compound of extracts. In this study, we used the clinical pathogen bacteria-resistant *Staphylococcus aureus*. TLC plates were carefully dried to remove all solvents before being overlaid with agar medium and left overnight at 37 °C. The areas of inhibition were compared with the R_f of the related spots on the reference TLC plate.

Anticancer Assay

The cytotoxic assay of the water fraction refers to Nurrachma *et al.* [14]. The water fraction was subjected to cell cytotoxic testing using two colorectal cancer cells (Caco-2 and DLD-1) and one normal cell (HEK293). The suspensions of colorectal cancer cells (Caco-2 and DLD-1) and normal cells (HEK293) were harvested at 70–80% confluency and centrifuged at 1200 rpm/25°C for 3 min. The cells were counted under a microscope using a haemocytometer and were then seeded into 96 well plates at 1 × 10⁵ cells per mL. Incubation was set to 5% CO₂ at 37°C for 24 h to allow for cell attachment. After that, the unattached cells were removed carefully. Cells were then treated with serial dilution of 100; 50; 25; 12.5; 6.25; and 3.125 µg/mL of sponge extracts and further incubated for 72 h. Then, MTT reagent (5 mg/mL) was added into each well and incubated at 5% CO₂, 37°C for 4 h. The cell media and MTT mixture were removed, and the purplish formazan crystals were dissolved in sodium dodecyl sulfate (SDS) and further incubated in a dark environment. The absorbance was read on the spectrophotometer at 570 nm wavelength. All experiments were performed in triplicate. The 50% inhibitory concentration (IC₅₀) of samples was calculated by statistical analysis. Cell viability was determined using the formula below:

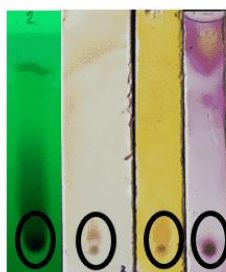
$$\text{Cell Viability} = (\text{Abs Treated cells} / \text{Abs Untreated cells}) \times 100\%$$

RESULTS AND DISCUSSION

Thin Layer Chromatography (TLC)

The visualization using UV light 254 nm (short wavelength) showed that the compound component of water fraction detectable band under UV light marked by the dark spot in R_f 0-0.6, indicating the presence of chromophores group of organic aromatic compounds that can absorb UV light 254 nm [15] (Figure 1a). The fraction positively reacted with cerium (IV) sulfate and showed a brown spot in the polar region with R_f 0-0.1, indicating which detects the presence of oxidizable organic compounds [16] (Figure 1b). Furthermore, the reaction with Dragendorff showed an orange spot at R_f 0-0.1, this indicates an alkaloid with tertiary N groups [17] (Figure 1c). Visualization using ninhydrin reagent shows a purplish stain at R_f 0, this indicates the presence of amino acids or amines [18] (Figure 1d). Based on the TLC, the water fraction showed

characteristics of alkaloid compounds, such as conjugated double bonds, tertiary N groups, and primary or secondary N groups. These results are consistent with previous studies that have found alkaloid compounds in sponges of the genus *Aaptos*, including 2-methoxy-3-oxoaaptamine and dimethyl(oxy)aaptamine (*Aaptos* sp. from Indonesia) [5], and demethylaaptamine (*Aaptos* sp. from Vietnam) [19]. A review by Nadar et al. [6] has also noted that sponges of the genus *Aaptos* may produce a variety of alkaloid compounds with therapeutic activities such as antibacterial and anticancer properties.



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

Figure 1. Visualization of TLC, (a) UV 254 nm, (b) $\text{Ce}(\text{SO}_4)_2$, (c) Dragendorff, (d) Ninhydrin.

Identification of Metabolite Using LCMS/MS

In this study, chromatogram data using the MassLynx program showed seven peaks approximately at retention time (RT) 1.11, 1.19, 4.02, 4.42, 4.62, 15.03, and 18.53 min as shown in Figure 2.

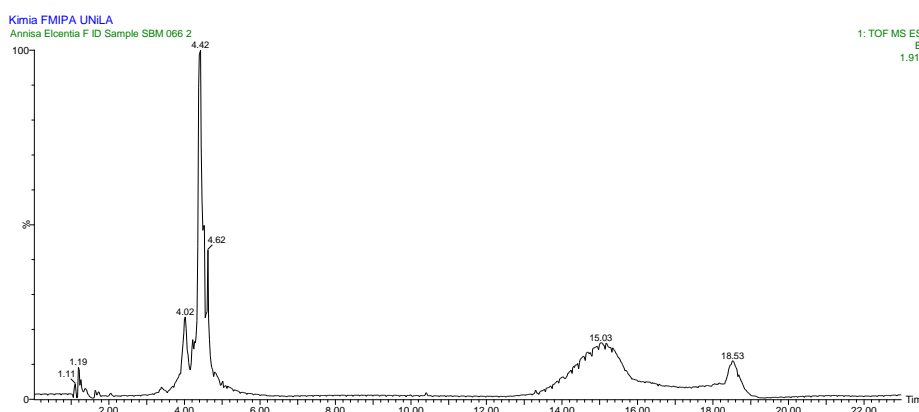


Figure 2. Chromatogram Water Fraction of *Aaptos* sp.

From these peaks, we can analyze the composition formula of the peak, respectively, as shown in Table 1.

Table 1. Identified Compound of The Each Peak

No	RT(Min)	Calc. Mass	Formula $[\text{M}+\text{H}]^+$	Fit. Conf (%)	Identified
1	1.11	151.0387	$\text{C}_3\text{H}_8\text{N}_4\text{OCl}$	65.97	Amino acid
2	1.19	217.1665	$\text{C}_9\text{H}_{21}\text{N}_4\text{O}_2$	99.99	Amino acid
3	4.02	229.0977	$\text{C}_{13}\text{H}_{13}\text{N}_2\text{O}_2$	99.98	Alkaloid
4	4.42	229.0977	$\text{C}_{13}\text{H}_{13}\text{N}_2\text{O}_2$	99.99	Alkaloid
5	4.62	229.0977	$\text{C}_{13}\text{H}_{13}\text{N}_2\text{O}_2$	99.94	Alkaloid
6	15.03	577.5169	$\text{C}_{37}\text{H}_{69}\text{O}_4$	81.77	Ester
7	18.53	577.5196	$\text{C}_{37}\text{H}_{69}\text{O}_4$	78.97	Ester

LC-MS provides data regarding molecular weight and structural characterization. It also helps in the detection and identification of impurities compounds [9]. Based on the elemental composition data in MassLynx, the retention time (RT), molecular weight, formula molecule, and % fit. configuration was obtained as shown in Table 1. The results of this study indicate that the detected compounds belong to the group of alkaloids, amino acids, and ester (Table 1). The compound groups (Table 1) are derived from the data

available in the PubChem. Alkaloid is the major component with >99% similarity, while the amino acid and ester were identified as minor metabolites. It is supported by our TLC result (Figure 1) which displayed the presence of alkaloids. According to He *et al.* [2], alkaloid plays an important role bioactivity of the sponge genus *Aaptos*.

It is known that there are 3 peaks (4.02, 4.42, and 4.62) that have identical positive ion ESI-MS-QTOF with the highest intensity at a retention time of 4.42 with a similarity of 99%. The metabolites were detected at a protonated molecular mass $[M+H]^+$ of m/z 229.0977 and formula molecules $C_{13}H_{13}N_2O_2$. The hydrogen atom from the data is reduced by one to get the suitable molecular formula in the database because the instrument LC-MS used has positive ions [20]. This means that 3 peaks actually have formula molecules $C_{13}H_{12}N_2O_2$ and have 9 degrees of unsaturation. Based on the MS and MS/MS data obtained, the molecular formula of the ion was speculated of $C_{13}H_{12}N_2O_2$, and then it was inferred of aaptamine. The results indicated that aaptamine is the major compound in this sample which was similar to the compound that was previously reported by Calcul *et al.* [21].

Identification of Metabolites Using ¹H-NMR

The characteristic ¹H-NMR chemical shifts are shown in Table 2.

Table 2. NMR spectroscopic data of aaptamine

Experiment ^a	Luyao <i>et al.</i> , 2021 ^b	δ_H (J in Hz)		No. atom H
		Calcul <i>et al.</i> , 2003 ^c	Nakamura <i>et al.</i> , 1982 ^d	
1,48 s				
1,52 s				
1,53 s				
1,64 s				
1,65 s				
3,04 d ($J=1,5$)				
3,28 s				
3,89 s	3,93 s	3,77 s	3,86 s	9-OMe
3,94 d ($J=7,0$)				
4,04 s	4,04 s	3,93 s	4,03 s	8-OMe
4,08 s				
4,88 s				
4,99 s				
5,02 s				
6,14 d ($J=7,0$)				
6,35 d ($J=7,5$)	6,34 d ($J=7,1$)	6,49 d ($J=7,0$)	6,52 (d, $J=6,5$)	3
6,80 d ($J=7,5$)				
6,88 d ($J=7,5$)	6,86 d ($J=7,2$)	6,78 d ($J=7,4$)	6,93 (d, $J=7,3$)	6
7,24 d ($J=6,5$)				
7,59 d ($J=7,5$)	7,23 d ($J=7,2$)	7,30 d ($J=7,1$)	7,45 (d, $J=7,3$)	5
7,78 d ($J=7,5$)	7,77 d ($J=7,1$)	7,80 d ($J=7,0$)	7,90 (d, $J=6,5$)	2
8,19 s	7,04 (s)	7,02 s	7,18 s	7

^a Measured at 500 MHz in MeOH-*d*₆; ^b Measured at 600 MHz in MeOH -*d*₄; ^c Measured at 400.13 MHz in DMSO-*d*₆.

The ¹H-NMR spectrum displayed the characteristic signals for two isolated methoxy protons 8-OCH₃ (δ_H 3.89, s) and 9-OCH₃ (δ_H 4.04, s), two sets of coupled protons at H-2 (δ_H 7.78, d, $J=7.5$ Hz) and H-3 (δ_H 6.35, d, $J=7.5$ Hz), and H-5 (δ_H 7.59, d, $J=7.5$ Hz) and H-6 (δ_H 6.88, d, $J=7.5$ Hz), and the isolated proton H-7 (δ_H 8.19, s). The spectroscopic data obtained matches the chemical shift values of the aaptamine compound (Figure 3), as reported on the Okinawan sponge *Aaptos aaptos* by Nakamura and coworkers [22], the Indonesian sponge *Xestospongia* sp. by Calcul and coworkers [21], and the Philippine sponge *Stylissa* sp. by Luyao and coworkers [23]. A slight chemical shift exists between the literature and experimental data due to the difference in solvents. In a solution, both the solute and solvent molecules have the ability to undergo polarization through electrostatic interactions. They can form specific bonds that involve electron transfer and

induce conformational changes [24]. The shimming process also plays a vital role in determining the accuracy of NMR spectrum results. If the shimming process is not optimal, it may lead to the broadening of the lines in the spectrum, resulting in reduced sharpness and clarity of the signals. This can make different interpretation and analysis, as pointed out by Noguchi in 2014 [25].

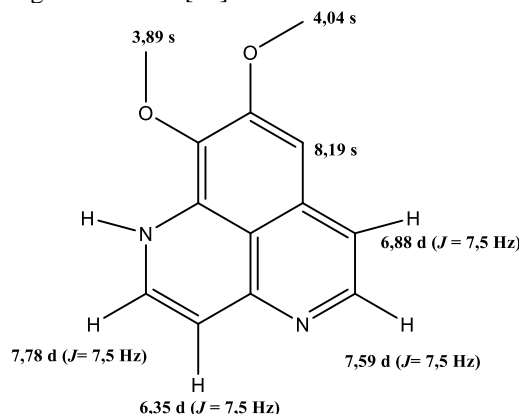


Figure 3. Structure Aaptamine

TLC Bioautography

The TLC bioautography of the sample revealed clear zones of bacterial growth inhibition for clinical-resistant *Staphylococcus aureus*. TLC bioautography is basically a simple measurement providing a response based on the antibacterial activity of its components through the spots they produce [26]. The clear zone with a diameter ± 5 mm resulted in the agar plate. It showed that the water fraction of *Aaptos* sp. can diffuse into the agar and inhibit the growth of the tested microorganism, suggesting that this compound has antibacterial activity as shown in Figure 4.

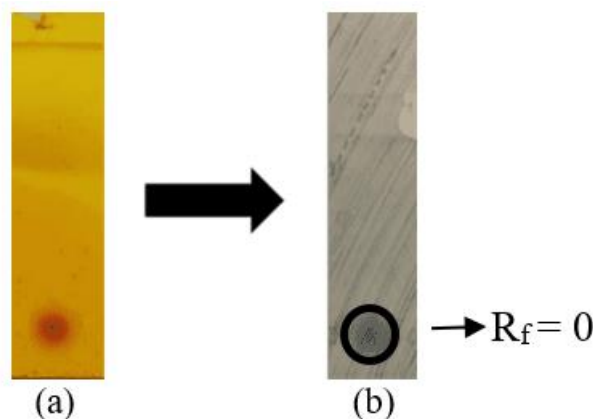


Figure 4. TLC Bioautography using mobile phase n-Hex:EtOAc (7:3) and stationary phase silica gel GF₂₅₄; (a) TLC Reference, (b) The clear zone of TLC Bioautography

Rajivgandhi *et al.* [27] was reported that aaptamine is promising antibacterial candidate against Gram-negative bacteria such as *Escherichia coli*, *Proteus mirabilis*, *Pseudomonas aeruginosa*, *Klebsiella pneumoniae*, and *Enterobacter* sp. It also potential for controlling *Vibrio harveyi* on tiger shrimp culture [28]. Aaptamine derivatives, 1,4-Dihexylaaptamine, can exhibit Gram-positive bacteria including *Bacillus cereus*, *Staphylococcus aureus* and *Micrococcus* sp [29]. The reports on aaptamine compounds as antibacterials are still very limited, especially reported on multi-drug resistance (MDR) antibacterial. Even though MDR is more complex to handle and presents a higher frequency of side effects [30]. To the best of our knowledge, this is the first time that aaptamine compounds have been tested on clinical bacteria-resistant and promising as antibacterial against clinical-resistant *Staphylococcus aureus*. The discovery of the antibacterial activity of this compound can be used as an alternative for treating infections caused by multi-drug resistant *S. aureus*.

MTT Cell Viability Assay

The compound has been evaluated for their cytotoxic effect on three cell lines including the colorectal adenocarcinoma cell line (Caco-2 and DLD-1) and normal cells (HEK293). The cytotoxicity test of water fraction can be seen in Table 3.

Table 3. IC₅₀ value of water fraction against cancer cells Caco-2, DLD-1, and HEK293 cancer cells.

No.	Cancer Cells	IC ₅₀ (µg/mL)	Category
1.	Caco-2	36.64	moderate
2.	DLD-1	28.97	moderate
3.	HEK293	3266.04	not active

Based on Table 2, water fraction has cytotoxic activity against colorectal adenocarcinoma cells (Caco-2 and DLD-1) and normal cells (HEK293), with an IC₅₀ value of 36.64; 28.97; and 3266.04 µg/mL, respectively. According to the National Cancer Institute (NCI), anticancer compound candidates are declared to have high cytotoxic activity if IC₅₀ <20 µg/mL, moderate cytotoxic activity if IC₅₀ ranges from 21-200 µg/mL, weak cytotoxic activity if IC₅₀ ranges from 201-500 µg/mL, and no cytotoxic activity if IC₅₀ > 500 µg/mL [31]. The results showed that water fraction has moderate cytotoxic activity as anticancer against Caco-2 (IC₅₀ 36.64 µg/mL) and DLD-1 (IC₅₀ 28.97 µg/mL) cells but has no cytotoxic activity against normal HEK293 cells (IC₅₀ 3266.04 µg/mL). The existence of higher cytotoxicity properties against cancer cells and low cytotoxicity against normal cells in anticancer compound candidates is highly preferred because it can minimize side effects [32]. Based on the data, the water fraction is known to have the potential as an anticancer agent for Caco-2 and DLD-1, without causing side effects to normal cells HEK293.

CONCLUSION

The rapid identification using LC-MS/MS combined with ¹H-NMR showed that the water extract of sponge *Aaptos* sp. collected from Bunaken National Park, North Sulawesi, Indonesia contained aaptamine. The water fraction can exhibit antibacterial activity against clinical bacteria-resistant *Staphylococcus aureus* with clear zone ±5 mm and has moderate cytotoxic activity as an anticancer against Caco-2 and DLD-1 colorectal cancer cells, but has no cytotoxic activity against normal HEK293 cells. These findings suggested that water fraction from *Aaptos* sp. contained aaptamine and can be used as an antibacterial as well as an anticancer agent. This research also demonstrated the importance of rapidly identifying bioactive compounds using LCMS/MS supported with MassLynx and PubChem databases to accelerate the discovery of compounds and hold promise strategy for future drug development.

ACKNOWLEDGEMENT

This research was financially supported by PEE (Pendanaan Ekspedisi dan Eksplorasi [Grant Code pee-2229172647]) and Prioritas Riset Nasional (PRN) also known as RIIM (Riset dan Inovasi untuk Indonesia Maju) [Grant code prn-012917423].

REFERENCES

- [1] Mehbub, M. F., Lei, J., Franco, C., & Zhang, W. Marine drugs, **12**(8), 4539-4577 (2014).
- [2] He, Q., Miao, S., Ni, N., Man, Y., & Gong, K. Natural Product Communications, **15**(9), 1934578X20951439 (2020).
- [3] Pham, C. D., Hartmann, R., Müller, W. E., De Voogd, N., Lai, D., & Proksch, P. Journal of natural products, **76**(1), 103-106 (2013).
- [4] Trang, D. T., Tai, B. H., Hang, D. T. T., Yen, P. H., Nhiem, N. X., & Kiem, P. V. Natural Product Research, **36**(19), 5022-5031 (2022).
- [5] Arai, M., Han, C., Yamano, Y., Setiawan, A., & Kobayashi, M. Journal of natural medicines, **68**, 372-376 (2014).
- [6] Nadar, V. M., Manivannan, S., Chinnaiyan, R., Govarthanan, M., & Ponnuchamy, K. Chemical Biology & Drug Design, **99**(1), 103-110 (2022).

- [7] Kurnianda, V., Wisna, U. J., & Musman, M. (2022). "Biodiversity of the Indonesian marine sponge genus *Aaptos* with molecular networking approach," in *E3S Web of Conferences*, (EDP Sciences), Vol. 339, p. 03002.
- [8] Sabdono, A. *Journal of Coastal Development*, **11**(2), 57-61 (2008).
- [9] Ma, X. *Molecules*, **27**(19), 6466 (2022).
- [10] Gaudêncio, S. P., Bayram, E., Lukić Bilela, L., Cueto, M., Díaz-Marrero, A. R., Haznedaroglu, B. Z., ... & Tasdemir, D. *Marine Drugs*, **21**(5), 308 (2023).
- [11] Ahmadi, P., Haruyama, T., Kobayashi, N., de Voogd, N. J., & Tanaka, J. *Chemical and Pharmaceutical Bulletin*, **65**(9), 874-877 (2017).
- [12] Setiawan, A., Lutfiah, R., Juliasih, N. L. G. R., Setiawan, W. A., Hendri, J., & Arai, M. *AACL Bioflux*, **15**(3), 1413-1421 (2022).
- [13] Chomnawang, M. T., Surassmo, S., Wongsariya, K., & Bunyaphatsara, N. *Fitoterapia*, **80**(2), 102-104 (2009).
- [14] Nurrachma, M. Y., Maran, G. G., Putri, N. B., Esti, Y. F., Hermawan, A., Meiyanto, E., & Jenie, R. I. *Indonesian Journal of Cancer Chemoprevention*, **11**(3), 103-114 (2020).
- [15] Pratiwi, R. A., & Nandiyanto, A. B. D. *Indonesian Journal of Educational Research and Technology*, **2**(1), 1-20 (2022).
- [16] Prashanth, K. N., & Basavaiah, K. *Proceedings of the National Academy of Sciences India Section A - Physical Sciences*, **84**(1): 27-35 (2014).
- [17] Raal, A., Meos, A., Hinrikus, T., Heinämäki, J., Române, E., Gudienė, V., & Nguyen, H. T. *Die Pharmazie-An International Journal of Pharmaceutical Sciences*, **75**(7), 299-306 (2020).
- [18] Das, S. Recent applications of ninhydrin in multicomponent reactions. *RSC advances*, **10**(32), 18875-18906 (2020).
- [19] Dyshlovoy, S. A., Fedorov, S. N., Shubina, L. K., Kuzmich, A. S., Bokemeyer, C., Keller-Von Amsberg, G., & Honecker, F. *BioMed Research International*, 2014: 1- 7 (2014).
- [20] Fitmawati, Khairunnisa, Resida, E., Emrizal, & Mustika Roza, R. *Journal of Physics: Conference Series*, **1351**(1), 012027 (2019).
- [21] Calcul, L., Longeon, A., Al Mourabit, A., Guyot, M., & Bourguet-Kondracki, M. L. *Tetrahedron*, **59**(34), 6539-6544 (2003).
- [22] Nakamura, H., Kobayashi, J. I., Ohizumi, Y., & Hirata, Y. *Tetrahedron Letters*, **23**(52), 5555-5558 (1982).
- [23] Luyao, H., Luesch, H., & Uy, M. *Molecules*, **26**(18), 5618 (2021).
- [24] Dračinský, M., & Bour, P. *Journal of Chemical Theory and Computation*, **6**(1), 288-299 (2010).
- [25] Noguchi, S. *Journal of Applied Physics*, **115**(16), (2014).
- [26] Wang, M., Zhang, Y., Wang, R., Wang, Z., Yang, B., & Kuang, H. *Molecules*, **26**(15), 4647 (2021).
- [27] Rajivgandhi, G., Ramachandran, G., & Manoharan, N. *Biocatalysis and Agricultural Biotechnology*, **17**, 628-637, (2019).
- [28] Rosmiati, R., Parenrengi, A., & Suryati, E. *Indonesian Aquaculture Journal*, **10**(1), 33-40, (2015).
- [29] Rashid, F. N. A. A., Asari, A., Mohamad, H., & Nor, S. M. *MAsian Journal of Chemistry*, **26**(20), 6903, (2014).
- [30] Guo, Y., Song, G., Sun, M., Wang, J., & Wang, Y. *Frontiers in cellular and infection microbiology*, **10**, 107, (2020).
- [31] Damasuri, A. R., & Sholikhah, E. N. *Indonesian Journal of Pharmacology and Therapy*, **1**(2), (2020).
- [32] Markman, J. L., Rekechenetskiy, A., Holler, E., & Ljubimova, J. Y. *Advanced drug delivery reviews*, **65**(13-14), 1866-1879 (2013).

Analysis of Limau Kuit Leaves Volatile Compounds Obtained by Soxhlet-N-Hexane Extraction with Gc-Ms

Azidi Irwan^{1,*}, Yuni Kurniawati¹, Kamilia Mustikasari¹

¹Department of Chemistry, Faculty of Mathematics and Natural Sciences, Lambung Mangkurat University, Jalan Jenderal Achmad Yani KM 35.8 Banjarbaru, 70714 South Kalimantan

*Corresponding Author: airwan@ulm.ac.id

ABSTRACT. This study aims to investigate the content of volatile compounds from the *n*-hexane extract of limau kuit leaves. Limau kuit is considered to be a typical citrus from South Kalimantan. The extraction was carried out using the soxhlet solvent extraction method, and then volatile compounds were determined using GC-MS. Limau kuit leaves were divided into four samples, each pre-treated before extractions conducted at 69°C. Those were fresh leaves with and without withering (by splashing of water); dried leaves passed a 20-mesh sieve, and dried leaves passed a 60-mesh sieve. The observations comprised the yield of each extraction, refractive index, and optical rotation values. The results are as follows: fresh leaves with withering (1.58%, 1.377, -0.21); fresh leaves without withering (0.23%, 1.373, -0.4); dried leaves passed 20-mesh (4.21%, 1.429, -); dried leaves passed 60-mesh (3.84%, 1.434, -), respectively. The results of GC-MS analysis in the same order as the previous sample showed that there were 29 compounds (the five major compounds were α -terpinene, *o*-cymene, thymol, spathulenol, and aromadendrene): 48 compounds (*o*-cymene, α -terpinene, thymol, 7-methyl-5-octen-4-one, and limonene-4-ol), 64 compounds (β -terpinene, (-)-spathulenol, *o*-cymene, caryophyllene, and germacrene D), and 70 compounds ((+)-spathulenol, germacrene D, caryophyllene, *o*-cymen-5-ol, and germacrene B). They overall indicated that α -terpinene, *o*-cymene, (+)-spathulenol, *b*-pinene, *a*-pinene, *o*-cymen-5-ol, limonen-4-ol, *g*-cadinene, *p*-mentha-1,4-dien-7-ol, and caryophyllene were the predominant compounds compared with the others.

Keywords: limau kuit, volatile compound, solvent extraction, *n*-hexane, soxhlet, GCMS

INTRODUCTION

Orange is one type of fruit that is most popular with Indonesian people. Each type of orange has different characteristics in shape, taste, and aroma. Oranges have been used by the community both for direct consumption as food and drinks and as food flavouring. One type of orange used as a flavouring for dishes such as soup, rice soup, and even the typical Banjar shrimp paste sauce is limau kuit (limau in the Banjar language means orange). Limau kuit is a type of orange typical of South Kalimantan province.

Limau kuit resemble the lime (Key Lime, Mexican Lime, *Citrus aurantifolia* (Christm.) Swingle) [1], and kaffir lime (*Citrus hystrix* D.C.). The similarity between the limau kuit and the Key Lime lies in the sour taste of the juice. The acid taste of the lime can be used as a flavouring substitute for limau kuit and when eating Banjarese soup because the lime is often used as a spice in the kitchen, as an ingredient in drinks, as a pharmacy, and as a herbal cough suppressant [2]. While the similarities between limau kuit and kaffir lime are presented in the morphology, fruit, stems, and leaves. Limau kuit and kaffir lime peels have wrinkled surfaces. Another difference is that kaffir lime leaves are similar to shape number 8, but the limau kuit leaves close to the stem are smaller, while kaffir lime leaves number 8 have the same size. Along the bark of the orange and kaffir lime plants, long and sharp thorns grow on the stems [3].

Based on literature studies on citrus research, research on limau kuit essential oil has not been found much. Therefore research on the chemical content refers to kaffir lime, key lime, and others that were taken from their leaves to produce essential oils. Various studies on orange leaves have varying concentrations of citronellal, linalool, and linalool acetate. The investigation results by Kasuan [4] using a steam distillation technique showed a high citronellal content from kaffir lime peels and 14 other compounds. The uniqueness of these compounds gives a different aroma to kaffir lime and lime. Limau kuit leaves are presumed to contain a unique composition of volatile compounds that give rise to a distinct smell. Finding new components to add to the uniqueness of the essential compounds from limau kuit leaves is even possible [5]

The right extraction method and solvent determine the quantity of essential oil yield. Previous research [6] showed that in the extraction of kaffir lime leaves using the Soxhlet technique using *n*-hexane and ethanol solvents, more citronellal was produced using *n*-hexane solvents than ethanol solvents. So in this study, the Soxhlet technique was used as an experiment because it uses less solvent, produces a more concentrated yield, and contains more active substances. While the use of *n*-hexane because this solvent is selective in dissolving substances and can extract large amounts of volatile substances [6].

In addition to the influence of the type of citrus, method of extraction, and type of solvent, previous research also involved the influence of other factors in extracting essential oils from citrus leaves. Another factor is the effect of treatment on the sample. The experiment involved four pretreatments: fresh leaves withered, leaves without withered, dry leaves passed a 15 mesh sieve, and dried leaves passed a 60 mesh sieve [7]. Wilted leaf samples can increase the yield of extracted essential oils. Based on previous research, preliminary treatment was carried out on kaffir lime leaf research referring to [7] because the study represented the treatment of fresh and dry leaves with four variations, namely fresh limau kuit leaves with and without withering, dried leaves passed through 20 mesh and 60 mesh sieves. The results of essential oil extraction were analyzed for yield, characteristics, and compounds. The characteristics include refractive index and optical rotation values. Meanwhile, to analyze the components of the extract used GC-MS. This research will observe the effect of various pretreatments on the characteristics of extracts on yield, optical rotation, and refractive index and determine their compounds.

MATERIALS AND METHODS

Materials

Limau kuit fresh leaves were collected from Sungai Tuan, Astambul, District of Banjar, and were used for all four experimental extractions. For the first two samples, before extraction, each of them was weighed at 150 grams and then sized to smaller by cutting about 0.5 cm in cross-section. These small cut leaves were then divided into two samples, the first one fresh leaves without withering and the second with withering. The other leaves were weighed 150 grams and air-dried for one week, then crushed and filtered with 20 and 60 mesh sizes. After extraction, the extract samples were stored at a chilled temperature before undergoing further quality analysis.

Samples Extraction

The limau kuit leaves from the four pretreatment variations were extracted using a Soxhlet extractor. The *n*-hexane solvent was added to the round flask. During the extraction process, the temperature of the *n*-hexane solvent was maintained not to exceed 69°C. The extraction process is completed by marking the colour of the solvent back to clear when it is evaporated (6-7 hours). The extraction result of each pretreatment obtained was essential oil extract mixed with *n*-hexane.

In the purification stage, this mixture of extract and *n*-hexane is separated using a vacuum rotary evaporator to eliminate the *n*-hexane. The same procedure was applied for evaporating whole samples to produce free-*n*-hexane viscous extracts.

Characterization of The Extract

The extract leaves were evaluated for characteristics, including yield, optical rotation, and refractive index. The yield values were obtained by calculating each final sample weight divided by the corresponding initial weight multiplied by 100%. Polarimeter AA-65 was used to measure the optical rotation of each sample. The polarimeter tube is filled with diluted extract (25% in *n*-hexane); no air bubbles should exist. The tube is placed into the polarimeter, and the value of the optical rotation of the oil is read on the scale panel. Optical rotation is expressed in degrees of circumference to close to 0.01°. The dextro angle is marked with a positive (+) sign, and the levo is negative (-). The refractive index measurement was carried out at a temperature of 27°C using the Refractometer Abbe 60. Both sides of the place to put the extract are cleaned with alcohol and allowed to dry. The tested extract is then dripped onto the surface and glued. The line settings are rotated to define the proper and centre boundaries of the light and dark surfaces. The scale reading can be done if the dividing line is right in the middle of the plane.

Analysis of Volatile Compounds

The volatile compounds of the extracts were analyzed using Hewlett Packard 6890 Gas Chromatograph with an MSD detector under the following operational conditions: injection volume of 1 μ L with split mode (ratio 20:1). The column used was HP-5MS, Agilent 19091S-433UI:0236716H with a stationary phase of 5% phenyl methyl siloxane with dimensions of 30 m in length x 250 μ m in diameter x

0.25 μ m in thickness. The carrier gas uses helium (mobile phase) at a constant 1 mL/min flow and a column pressure of 10.45 psi. The column temperature was programmed from 100°C to 325°C with an increase in temperature of (i) 10°C per minute to 100°C and maintained for 10 minutes, (ii) 10°C per minute to 150°C and maintained for 4 minutes. During the analysis, the injector temperature is programmed to a constant temperature of 325°C.

The identification of the extracted chemical compounds was referred to the mass spectral library of the instrument, NIST02.L, and Wiley275.L database.

RESULTS AND DISCUSSION

Yield and physical analysis of limau kuit leaves extract

Limau kuit leaves have been extracted through three stages: pretreatment, soxhlet extraction, and solvent separation (purification), followed by analyzing essential oil characteristics and compositions. The results of the leaf extractions with the pretreatment are shown in Figure 1.

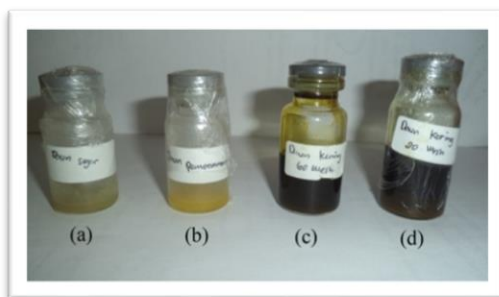


Figure 1. The limau kuit leaves *n*-hexane extracts from (a) fresh leaves without withering, (b) fresh leaves with withering, (c) dried leaves passed 20 mesh sieving, and (d) dried leaves passed 60 mesh sieving.

Based on Figure 1, pretreatment of fresh leaves without (a) and with (b) withering produces essential oils with yellowish-green and slightly greenish colours. Pretreatment of dried leaves through 20 (c) and 60 mesh (d) sieves had dark green extract to black. The cause of the difference in the colour of the extract results may be influenced by the colour factor of the leaves during the pretreatment. The solvent can absorb the compounds present in the leaves and the dyes, which aligns with research [8] regarding the extraction of essential oils from kaffir lime leaves, producing fresh essential oils that are greenish-yellow. Based on comparing the colours extracted from limau kuit and kaffir lime leaves, it is presumed that the two leaves have relatively similar essential oil compounds.

Table 1. The values of extract yields, refractive index, and optical rotation.

Pretreatment	Yield (%)	Refractive Index	Optical Rotation
Fresh leaves without withering (a)	0.23	1.374	-0,41°
Fresh leaves with withering (b)	1.58	1.377	-0,12°
Air-dried leaves passed 20 mesh sieving (c)	4.21	1.429	unread
Air-dried leaves passed 60 mesh sieving (d)	3.84	1.434	unread

The yield of essential oil extraction with four various pretreatments shows in Table 1. The result of the air-dried samples was relatively higher (pretreatment of (c) 4.21%, and (d) 3.84%) than fresh leaves (pretreatment of (a) 0.23% and (b) 1.58%). Fresh leaves are thought to make it difficult for solvents to penetrate the cells or oil gland pockets in the leaves due to the high water content. When solvent evaporation occurs during extraction, drying, and shrinkage of oily materials more easily penetrated by the liquid, the solvent in the gas phase will speed up the extraction process and reduce the water content contained in the material [9]. Table 2 compares the yield and refractive index of several other studies related to kaffir lime leaves.

Table 2. The value of yields and refractive index of the kaffir lime leaves from other research

Parameters	Sukardi, <i>et al.</i> (2014)		Maya-sari <i>et al.</i> (2013)		Munawarah & Handayani (2010)		Khasanah <i>et al.</i> (2015)			
	Fresh	Dried	Fresh	Dried	Dried (<i>n</i> -hex)	Dried (EtOH)	F.L.*	FLW	DL15	DL60
Yield (%)	0,23	1,58	0,78	10,50	13,39	-	-	-	-	-
Refractive Index	1,446	1,451	-	-	-	1,451	1,450	1,453	1,452	-

Ket: FL: Fresh Leaves, FLW: Fresh leaves with withering, DL15: Dried Leaves 15 mesh, DL60: Dried Leaves 60 mesh.

Table 1 shows the refractive index values of limau kuit leaf extract pretreatment (a), (b), (c), and (d) were 1.374; 1.377; 1.429, and 1.434, respectively. The values of the refractive index of the four pretreatments showed a slight difference among themselves (Table 1) and between the kaffir limes oil (Table 2), although they tend to be smaller [9]. Based on the data in Table 1 and Table 2 shows that the refractive index value of limau kuit extract is relatively similar to kaffir lime leaves oil, so it can be assumed that the purity of limau kuit leaf extract is excellent.

Chemical analysis of the extracts

The purpose of analysis using the GC-MS is to determine the compounds that characterize the limau kuit extract. This instrument is a combination of GC and MS functional tools. Figures 2, 3, 4, and 5 below present the following chromatograms of GC analysis results.

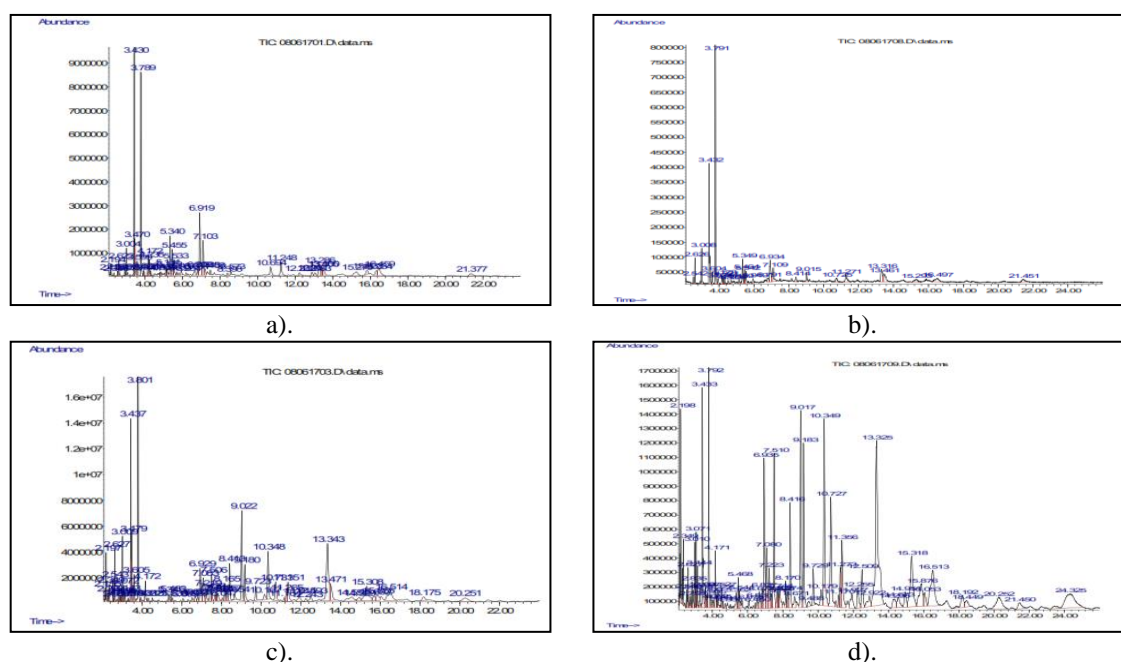


Figure 2. Typical GC-MS chromatograms showing the chemical analysis of the leaf extracts from each pretreatment: a). fresh without withering. The MS datasheet showed forty-eight contained compounds. b). fresh with withering (twenty-nine compounds), c). air-dried passed 20 mesh sieving (sixty-four compounds), and d). air-dried passed 60 mesh sieving (seventy compounds).

Based on Figures 2, 3, 4, and 5, each chromatogram shows the number of compounds from pretreatment (a) of as many as 48 and pretreatment (b) of 29 compounds. While from the pretreatment (c) and (d) were 64 and 70 compounds, respectively. The chromatogram shows the overlapping retention times of the GC-MS analysis results. The possibility is that there are similar molecular weights of the terpene groups, which are structural or geometric isomers of one another. As is known, the volatile compounds that make up essential oils are mostly terpene group compounds, which are dominant from the monoterpene (C10) and sesquiterpene (C15) groups or from their oxygenated groups (terpenoids).

The most significant amount in citrus essential oil is the terpene group with the basic structure of isoprene (C₅H₈, simplified C₅). Terpenes are formed from 2 isoprene units which are then called terpene units. The more prominent terpenes are multiples of 5 of the terpene units, so they are known as C₁₀, monoterpenes; C₁₅, sesquiterpenes; C₂₀, diterpenes; C₃₀, triterpenes; and C₄₀ tetraterpene, formulated with (C₅H₈)_n [10], [11]. Monoterpene, C₁₀, is the most abundant terpene in essential oils, followed by sesquiterpenes. Although C₂₀, C₃₀, and C₄₀ terpenes can be found, the concentrations are minimal.

Table 3 shows the distribution of volatile compounds obtained by GC-MS analysis from the pretreatment of fresh leaves without withering, with withering, air-dried leaves passed through a 20 mesh sieve, and air-dried leaves passed through a 60 mesh sieve. The data in the table is arranged based on the appearance of certain compounds in each pretreatment. These compound groups are marked with coloured shading. The yellow colour indicates that these compounds appear in all extracts. The red colour indicates the appearance of the three pre-treated extracts. The brown colour appeared in the two pretreatments, and the grey colour appeared only in one type of extract. Data is also arranged based on relative levels indicated by the area on the chromatogram of each extract. From top to bottom shows the levels from big to small.

Table 3. Comparative composition (%) of the *n*-hexane limau kuit leaves extracts of four applied pretreatments

No	Compound	a*	b*	c*	d*	Molecular Formula
1	<i>o</i> -cymene	16.08	14.29	7.26	3.39	C ₁₀ H ₁₂
2	(+)-spathulenol	3.07	4.99	13.02	12.22	C ₁₅ H ₂₄ O
3	<i>b</i> -pinene	1.76	3.89	2.39	0.86	C ₁₀ H ₁₆
4	<i>a</i> -pinene	0.97	2.06	1.68	0.56	C ₁₀ H ₁₆
5	<i>o</i> -cymen-5-ol	10.63	-	3.08	4.54	C ₁₀ H ₁₄ O
6	limonen-4-ol	4.27	4.03	0.49	-	C ₁₀ H ₁₆ O
7	<i>g</i> -cadinene	2.68	-	0.73	0.80	C ₁₅ H ₂₄
8	<i>p</i> -cymene	1.87	1.44	1.18	-	C ₁₀ H ₁₂
9	<i>tau</i> -cadinol	1.22	-	1.25	2.13	C ₁₅ H ₂₆ O
10	<i>cis</i> - <i>b</i> -ocymene	1.00	-	1.07	0.33	C ₁₀ H ₁₆
11	carvacrol	0.97	-	2.37	1.99	C ₁₀ H ₁₄ O
12	<i>cis</i> -sabinene hydrat	0.49	1.24	0.20	-	C ₁₀ H ₁₈ O
13	<i>a</i> -thujene	0.28	0.49	0.26	-	C ₁₀ H ₁₆
14	<i>p</i> -xylene	0.21	-	1.39	3.18	C ₈ H ₁₀
15	<i>a</i> -amorphene	-	3.19	0.30	0.33	C ₁₅ H ₂₄
16	caryophyllene	-	2.11	6.29	4.56	C ₁₅ H ₂₄
17	<i>b</i> -elemene	-	1.10	2.81	2.54	C ₁₅ H ₂₄
18	phenylethyl alcohol	-	0.90	0.16	0.37	C ₈ H ₁₀ O
19	<i>p</i> -mentha-1,4-dien-7-ol	14.93	-	-	3.19	C ₁₀ H ₁₆ O
20	copaene	3.06	-	0.93	-	C ₁₅ H ₂₄
21	<i>p</i> -cymen-4-ol	2.74	2.10	-	-	C ₁₀ H ₁₄ O
22	D-limonene	2.56	-	2.72	-	C ₁₀ H ₁₆
23	<i>a</i> -terpineol	1.95	2.24	-	-	C ₁₀ H ₁₈ O
24	<i>a</i> -terpinolene	1.42	0.80	-	-	C ₁₀ H ₁₆
25	caryophyllene oxide	1.30	-	-	0.70	C ₁₅ H ₂₄ O
26	<i>cis</i> -sabinene hydrat	0.90	1.24	-	-	C ₁₀ H ₁₈ O
27	elemol	0.88	-	0.32	-	C ₁₅ H ₂₆ O
28	<i>cis</i> -linalool oxide	0.40	-	0.25	-	C ₁₀ H ₁₈ O ₂
29	ethylbenzene	0.15	-	0.26	-	C ₈ H ₁₀
30	<i>g</i> -terpinene	-	24.53**	10.05	-	C ₁₀ H ₁₆
31	<i>cis</i> - <i>b</i> -terpineol	-	1.14	-	0.62	C ₁₀ H ₁₈ O
32	isospathulenol	-	-	5.05	1.91	C ₁₅ H ₂₄ O
33	Germacone D	-	-	4.09	5.75	C ₁₅ H ₂₄
34	<i>tau</i> -muurolol	-	-	3.13	4.01	C ₁₅ H ₂₆ O
35	<i>g</i> -elemene	-	-	2.63	1.50	C ₁₅ H ₂₄
36	decane	-	-	0.57	0.78	C ₁₀ H ₂₂
37	mesitylene	-	-	0.50	0.86	C ₉ H ₁₂
38	tridecane	-	-	0.40	0.47	C ₁₃ H ₂₈
39	dodecane	-	-	0.39	0.52	C ₁₂ H ₂₆
40	4-methyldecane	-	-	0.14	0.35	C ₁₁ H ₂₄
41	7-methyl-5-octen-4-one,	5.34	-	-	-	C ₉ H ₁₆ O
42	<i>p</i> -1,8-menthadien-4-ol/limonen-4-ol	4.27	-	-	-	C ₁₀ H ₁₆ O

No	Compound	a*	b*	c*	d*	Molecular Formula
43	epizonarene	3.14	-	-	-	C ₁₅ H ₂₄
44	<i>p</i> -cymene	2.49	-	-	-	C ₁₀ H ₁₂
45	(+)-valencene	2.38	-	-	-	C ₁₅ H ₂₄
46	1,4-dimethyl-7-(1-hydroxy-1-methylethyl)azulene	2.22	-	-	-	C ₁₅ H ₁₈ O
47	β -eudesmol	1.40	-	-	-	C ₁₅ H ₂₆ O
48	2,5-dimethyl-3-hexyne-2,5-diol	1.30	-	-	-	C ₈ H ₁₄ O ₂
49	oplopenone	1.21	-	-	-	C ₁₅ H ₂₄ O
50	<i>trans</i> -carveol	1.07	-	-	-	C ₁₀ H ₁₆ O
51	2-octen-4-ol	0.84	-	-	-	C ₈ H ₁₆ O
52	(3 <i>E</i> ,7 <i>E</i>)-4,8,12-trimethyltrideca-1,3,7,11-tetraene	0.83	-	-	-	C ₁₆ H ₂₆
53	4,5-epoxy-1-isopropyl-4-methyl-1-cyclohexene	0.73	-	-	-	C ₁₀ H ₁₆ O
54	(<i>E,E</i>)-farnesal	0.66	-	-	-	C ₁₅ H ₂₄ O
55	(<i>E</i>)-4,8-dimethyl-3,8-nonadien-2-on	0.59	-	-	-	C ₁₁ H ₁₈ O
56	1-(2-furanyl)-ethanone	0.45	-	-	-	C ₆ H ₆ O
57	<i>E,E</i> -2,6-dimethyl-3,5,7-octatriene-2-ol	0.39	-	-	-	C ₁₀ H ₁₆ O
58	2,3-dihydro-2-methyl-benzofuran	0.29	-	-	-	C ₉ H ₁₀ O
59	<i>b</i> -sabinene	0.26	-	-	-	C ₁₀ H ₁₆
60	ascaridole	0.21	-	-	-	C ₁₀ H ₁₆ O ₂
61	(2 <i>R</i> ,5 <i>S</i>)-2-methyl-5-(prop-1-en-2-yl)-2-vinyltetrahydrofuran	0.19	-	-	-	C ₁₀ H ₁₆ O
62	thymol	-	5.94	-	-	C ₁₀ H ₁₄ O
63	aromadendrene	-	4.42	-	-	C ₁₅ H ₂₄
64	1,8-mentadien-4-ol/limonen-4-ol	-	4.03	-	-	C ₁₀ H ₁₆ O
65	3 <i>a</i> ,6 <i>a</i> -dimethyl-1,2,3,4,5,6-hexahydropentalene	-	4.03	-	-	C ₂₀ H ₂₄ N ₂ O ₂ P
66	columbin	-	3.00	-	-	C ₂₀ H ₂₂ O ₆
67	2,4-dimethylamphetamine	-	2.66	-	-	C ₁₁ H ₁₇ N
68	3-methyl-4(3H)-pyrimidinone	-	2.13	-	-	C ₅ H ₆ N ₂ O
69	<i>a</i> -muurolene	-	1.64	-	-	C ₁₅ H ₂₄
70	<i>a</i> -ocymene	-	1.55	-	-	C ₁₀ H ₁₆
71	2-bromoethanol	-	1.24	-	-	C ₂ H ₅ BrO
72	prehnitene	-	1.14	-	-	C ₁₀ H ₁₄
73	1,8-dimethyl-4-(1-methylethyl)-spiro[4.5]decan-7-on	-	0.81	-	-	C ₁₅ H ₂₆ O
74	dimethyl octadecanadioic	-	0.64	-	-	C ₂₀ H ₄₀ O ₂
75	4-dienecadina-1(10)	-	-	2.29	-	C ₁₅ H ₂₄
76	<i>b</i> -gurjunene	-	-	2.24	-	C ₁₅ H ₂₄
77	<i>a</i> -farnesene	-	-	1.74	-	C ₁₅ H ₂₄
78	4-methyl-7-(1-hydroxy-1-methylethyl)azulene-1-carbaldehyde	-	-	1.67	-	C ₁₅ H ₁₈ O
79	D,L- <i>p</i> -fluorophenylalanine	-	-	1.18	-	C ₉ H ₁₀ FNO ₂
80	<i>Z,Z,Z</i> -1,5,9,9-tetramethyl-1,4,7-cycloundecatriene	-	-	1.11	-	C ₁₅ H ₂₄
81	1-formyl-2,2,6-trimethyl-3-cis-(3-methylbut-2-enyl)-5-cyclohexene	-	-	1.11	-	C ₁₅ H ₂₄ O
82	<i>trans</i> -linalool oxide	-	-	0.75	-	C ₁₀ H ₁₈ O ₂
83	(+)-cyclosativene ylangene	-	-	0.19	-	C ₁₅ H ₂₄
84	<i>cis</i> - <i>p</i> -mentha-2,8-dien-1-ol	-	-	0.37	-	C ₁₀ H ₁₆ O
85	2,5-diethylfuran	-	-	0.54	-	C ₈ H ₁₂ O
86	camphor	-	-	0.56	-	C ₁₀ H ₁₆ O
87	5-Ethylcyclopent-1-ene-1-carboxylic acid	-	-	0.74	-	C ₈ H ₁₂ O ₂
88	(1 <i>sr</i> ,2 <i>sr</i> ,1' <i>sr</i>)-1-(1-methoxythyl)-2-vinylcyclobutane	-	-	0.60	-	C ₉ H ₁₆ O
89	limonen-1,2-diol	-	-	0.84	-	C ₁₀ H ₁₈ O ₂
90	isolekene	-	-	0.80	-	C ₁₅ H ₂₄

No	Compound	a*	b*	c*	d*	Molecular Formula
91	calamenene	-	-	0.74	-	C ₁₅ H ₂₂
92	<i>o</i> -ethyltoluene	-	-	0.35	-	C ₉ H ₁₂
93	nonane	-	-	0.21	-	C ₉ H ₂₀
94	(2 <i>R</i> ,3 <i>S</i>)-rel-2,3-dimethylcyclohexanone	-	-	0.11	-	C ₈ H ₁₄ O
95	myristic acid	-	-	-	4.13	C ₁₄ H ₂₈ O ₂
96	<i>d</i> -elemene	-	-	-	3.20	C ₁₅ H ₂₄
97	(1 <i>S</i> ,2 <i>S</i>)-(+)-2-amino-3-methoxy-1-phenyl-1-propanol	-	-	-	1.27	C ₁₀ H ₁₅ NO ₂
98	undecane	-	-	-	0.94	C ₁₁ H ₂₄
99	<i>a</i> -cubenene	-	-	-	0.93	C ₁₅ H ₂₄
100	3-methylbutanal methylformylhydrazone	-	-	-	0.72	C ₇ H ₁₄ N ₂ O
101	2,5-dimethyl-2,5-diol-3-hexyne	-	-	-	0.56	C ₈ H ₁₄ O ₂
102	tricyclo[4.3.1.1(3,8)]undecane	-	-	-	0.55	C ₁₁ H ₁₈
103	(<i>S</i>)-cis-verbenol	-	-	-	0.53	C ₁₀ H ₁₆ O
104	spiro[5.5]undecan-1-one	-	-	-	0.48	C ₁₁ H ₁₈ O
105	terpinene-4-ol	-	-	-	0.48	C ₁₀ H ₁₈ O
106	(<i>Z</i>)-methyl ester-3-methoxy-2-methyl-2-butyric acid	-	-	-	0.45	C ₇ H ₁₂ O
107	1-ethyl-3methylbenzene	-	-	-	0.46	C ₉ H ₁₂
108	(1-methylpropyl) benzene	-	-	-	0.38	C ₁₀ H ₁₄
109	6-methyloctadecane	-	-	-	0.29	C ₁₉ H ₄ O
110	<i>trans-b</i> -ocymene	-	-	-	0.24	C ₁₀ H ₁₆
111	1-methyl-2-phenylcyclopropane	-	-	-	0.22	C ₁₀ H ₁₂
112	3-ethyl- <i>o</i> -xylene	-	-	-	0.19	C ₁₀ H ₁₄
113	3-carene	-	-	-	0.17	C ₁₀ H ₁₆
114	crocetane	-	-	-	0.11	C ₂₀ H ₄₂
115	durene	-	-	-	0.11	C ₁₀ H ₁₄

* a = Without withering, b = with withering, c = passed of 20 mesh sieving, d = passed of 60 mesh sieving

** highest content

The total number of compounds from the four pretreatments was 115, distributed according to their content. The distribution of the number of C atoms ranges from 10 C atoms to 20 C atoms (C₁₀-C₂₀). Table 3 shows compounds typical and common as constituents of essential oils, but some are also unusual and even less well-known hydrocarbons. It was possibly due to the ability of *n*-hexane to dissolve many compounds from plant tissues, especially limau kuit leaves.

CONCLUSION

The extraction by *n*-hexane solvent was performed with four pretreatments on limau kuit leaves before soxhlet extraction. The results showed that the pretreatment of dried leaves passing through a 20 mesh sieve had the highest yield value of 4.21%, while other treatments, namely fresh leaves with and without withering and dried leaves that passed a 60 mesh sieve, were 1.58% in a row; 0.23%; and 3.84%, respectively. The values of the refractive index analysis of fresh leaves with and without withering, dry leaves that passed the 20 mesh and 60 mesh sieves were 1.377; 1.373; 1.429; and 1.434, respectively. The optical rotation value analysis results of fresh leaves with and without withering were -0.21 and -0.41. In contrast, the treatment of dried leaves passed the 20 mesh and 60 mesh sieves the values were unreadable.

The GCMS data analysis showed that the extracts of the leaves without withering had 48, the leaves with had 29, the dried leaves passed 20 mesh sieving had 64, and the dried leaves passed 60 mesh sieving had 70 compounds. From the whole compounds, it was found that α -terpinene, *o*-cymene, (+)-spathulenol, *b*-pinene, *a*-pinene, *o*-cymen-5-ol, limonen-4-ol, *g*-cadinene, *p*-mentha-1,4-dien-7-ol, and caryophyllene were the predominant compounds compared with the others.

REFERENCES

- [1]. UCR, *Givaudan Citrus Variety Collection at UCR-Mexican Lime*, <https://citrusvariety.ucr.edu/crc1710>, (2023).
- [2]. B. Sarwono, *Jeruk dan Kerabat*, Penebar Swadaya, Jakarta, (1986).
- [3]. Kompas, *Beda Jeruk Nipis, Jeruk Lemon, Jeruk Purut, dan Jeruk limau (limo)*, <https://agri.kompas.com/read/2022/10/10/142000884/perbedaan-jeruk-nipis-jeruk-purut-jeruk-limau-dan-jeruk-lemon?page=all>. (2020).
- [4]. N. Kasuan, Z. Muhammad, Z. Yusoff, M.H.F. Rahman, M.N. Taib, & Z.A. Haiyee, Extraction of *Citrus hystrix* D.C. (Kaffir lime) Essential Oil Using Automated Steam Distillation Process: Analysis of Volatile Compound. *Malaysian Journal of Analytical Science*, **17**, 359-369 (2013).
- [5]. D. Mayasari, A. Jayuska, & M. A. Wibowo, Pengaruh Variasi Waktu dan Ukuran Sampel terhadap Senyawa Minyak Atsiri dari Daun Jeruk Purut (*Citrus hystrix* DC.), *JKK.*, **2**, 74-77 (2015).
- [6]. S. Munawaroh, & P.A. Handayani, Ekstraksi Minyak Daun Jeruk Purut (*Citrus hystrix* DC) dengan Pelarut Etanol dan *n*-Heksana, *Jurnal Kompetensi Teknik*, **2**, 73-78 (2010).
- [7]. L.U. Khasanah, K.R. Utami, & Y.M. Aji, Pengaruh Perlakuan Pendahuluan Terhadap Karakteristik Mutu Minyak Atsiri Daun Jeruk Purut (*Citrus hystrix* DC), *Jurnal Aplikasi Teknologi Pangan*, **4**, 48-55 (2015).
- [8]. F.K. Hidayat, Ekstraksi Minyak Atsiri dari Daun Jeruk Purut (*Citrus hystrix* DC) pada Skala Pilot-Plant, *Thesis*, Institut Pertanian Bogor, 1641 (1999).
- [9]. S. Ketaren, *Pengantar Tekhnologi Minyak dan Lemak Pangan*, Universitas Indonesia, Jakarta, 77-80 (1985).
- [10]. M. Sawamura, *Citrus Essential Oils: Flavor and Fragrance*. In *Citrus Essential Oils: Flavor and Fragrance*. John Wiley & Sons Inc., Hoboken, New Jersey (2010)
- [11]. P. Tongnuanchan & S. Benjakul, Essential Oils: Extraction, Bioactivities, and Their Uses for Food Preservation (Review), *Journal of Food Science*, **79**(7), R:1231-1249 (2014). DOI: 10.1111/1750-3841.12492.

Ruthenium Complex and Methyl Dendrimer Compounds as Artificial Photosynthesis. First Principle Study for Renewable Energy

Rahmat Gunawan^{*(1)}, Fahmi Fadillah⁽¹⁾, Desmond Kristian⁽¹⁾, Chairul Saleh⁽²⁾, Erwin⁽²⁾

⁽¹⁾ *Lab. of Physical Chemistry, Dept. of Chemistry, Faculty Math and Natural Sciences, Mulawarman University, Gunung Kelua, Samarinda, 75123 East Kalimantan, Indonesia*

⁽²⁾ *Lab. of Organic Chemistry, Dept. of Chemistry, Faculty Math and Natural Sciences, Mulawarman University, Gunung Kelua, Samarinda, 75123 East Kalimantan, Indonesia,*

Corresponding author: gunawan@fmipa.unmul.ac.id

Abstract. A calculation of the potential application of a methyl dendrimer as a light-absorbing molecule in an artificial photosynthetic system using density functional theory method was completed. The objectives of this research were to determine the ability of polyphenylene dendrimer molecules in the electron transfer and to study a chlorophyll-like polyphenylene methyl dendrimer in terms of its ability to harvest light. For a ruthenium complex after a stable structure was obtained, its UV-Vis absorbance was tested. Absorption was found at wavelengths of 1445.54, 1188.49, 943.09, 71.20, and 67.48 nm. Based on its HOMO and LUMO positions, the ruthenium complex was also suggested to be capable of electron transfers. For the optimized polyphenylene methyl dendrimer structure was found to have -8390.9026 J of energy with RMS value of 0.0002 J. Electron transfer activity occurs around where HOMOs and LUMOs are found in the polyphenylene dendrimer at -1.0453x10⁻⁷/- 1.9022 nm with 10-18 J of energy and wavelength absorption within UV spectrum range, suggesting the ability to harvest light similar to that of chlorophyll.

Keywords: -

INTRODUCTION

Hydrogen is widely used in conventional internal combustion engines or fuel cells that convert chemical energy directly into electrical energy without combustion [1]. Hydrogen production process requires either conversion of natural gas by steam or, as a more environmentally-friendly alternative, electrolysis of water into hydrogen and oxygen. The former results in carbon dioxide as the by product [2].

As far as ecological safety is concerned, water electrolysis that breaks water down into hydrogen and oxygen is the best choice. This method is built upon the principle of photosynthesis. Researchers commonly refer to this method as artificial photosynthesis. Artificial photosynthesis is an artificial chemical process mimicking photosynthesis, the process that produces carbohydrate and oxygen out of sunlight, water and carbon dioxide. A photocatalytic decomposition of water produces protons as one of the intermediates, which finally transforms into hydrogen molecules, and oxygen molecules as well [3].

Chlorophyll and photosynthesis are very closely related since chlorophyll is the primary agent of photosynthesis that enables plants to absorb energy from sunlight. In photosynthesis, chlorophyll vitally serves two primary functions. Most chlorophyll molecules (up to several hundreds of molecules per photosystem) functions as means of absorbing sunlight and transferring light energy through resonance energy transfer to a specific pair of chlorophyll molecules at the photosystem reaction center. The ruthenium complex and Methyl Dendrimer is a class of tree-shaped molecules that possess a unique ability to harvest light. The light harvesting process is inspired by a naturally-occurring mechanism, i.e. the natural photosynthetic mechanism. In natural photosynthesis, chlorophyll serves as the light harvester; a function that is parallel to that of dendrimers in artificial photosyntheses [4]. In order to further study the chemical processes involved in photosynthesis, it is necessary to conduct a computational research to further investigate the quantum mechanical aspects of artificial photosynthesis with The ruthenium complex and a polyphenylene methyl dendrimer as the chlorophyll substitute. The research shall be meant to study the molecular orbitals, optimize the structure and calculate energy of the optimized structure of the dendrimer so as to provide a clear picture of its properties that may serve as a preliminary data for further laboratory-scale experiments [5].

MATERIALS AND METHODS

The methods included a series of theoretical studies to determine The ruthenium complex and the polyphenylene dendrimers properties as a substitute for chlorophyll in artificial photosynthesis:

Calculations Methods

Light-harvesting compound and an agent of electron transfer by calculating its molecular orbitals. Geometric structure and energy graphic along with energy levels of both the highest occupied and the lowest unoccupied molecular orbitals (HOMO and LUMO) energies as the molecular descriptors [6].

Electronic Structure Methods

DFT-based calculations were accomplished using GAMESS while the recapitulation of electronic structure calculations was processed using linear regression analysis [7].

RESULTS AND DISCUSSION

All structure optimizations in this experiment used universal force field (UFF) calculation model. UFF was chosen because it takes all types of force field found in the periodic table into account. Optimization calculation model of UFF is as follows. The potential energy of any given geometry calculated for a particular molecule is written as a superposition of various interactions between two, three, and four atoms

Molecular Orbital Analysis for The Ruthenium Complex

From the visualization data above, it is shown that the HOMOs are located about the center of this molecule while the LUMOs near the double-core complex of manganese. The part of this compound that has HOMOs, which means that its orbitals are fully occupied by electrons, shall undergo intervention by ruthenium which will subsequently lead to absorption of energy from sunlight. This absorption will then cause an excess of electrons in the part of the compound so that this excess will be passed on to the acceptor part of the compound. Upon passing on of electron excess, the then HOMOs shall now transform into LUMOs. The other LUMOs, which have been so to begin with, shall then extract electrons from water the compound is being exposed to. This will cause the then LUMOs transform into HOMOs and thereby ready to donate electrons to the center part of the compound which now has LUMOs from the aforementioned process.

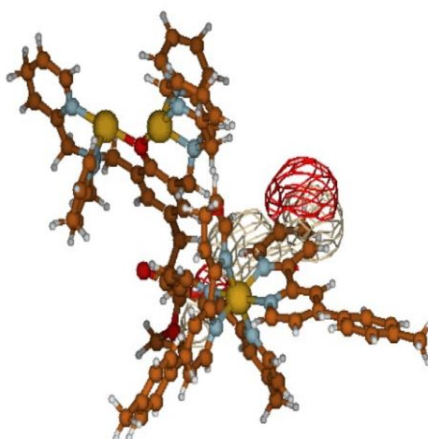


Figure 1. Molecular orbital of the ruthenium complex.

These processes are repeated to produce O₂ and H₂ as by products. To predict rutheniums efficacy in harvesting sunlight, the energy (RUNTYP=ENERGY) and UV/Vis absorption (CITYP=CIS) calculations were run for this element using RHF method. Fig. 1 is one of the results of UV/Vis absorption calculations for elemental ruthenium which displays its respective energies, excited states, and wavelengths. Wavelengths shown represent the range of wavelengths ruthenium is able to absorb in its excited state, i.e. 1445-67 nm. Chlorophyll has an absorption range at 750-380 nm of wavelength; the visible wavelength range. This

indicates that ruthenium is able to harvest light as chlorophyll does because the visible range is within ruthenium absorption range. This is strengthened by HOMO-LUMO arrangement in the ruthenium complex which is well-placed for electron transfers. In short, the structure of this ruthenium complex is effective enough for absorbing light and transferring electrons although not as effective as chlorophyll. Judging from the rate of technological advancement today, it is possible to develop this compound further in order to make it work better than or at least comparable to natural chlorophyll.

Molecular Orbital Analysis for the Polyphenylene Metyl Dendrimer

After the compound's structure was optimized and its energy was determined, the next step was to study the molecular orbitals of both the polyphenylene dendrimer and chlorophyll with the help of Gmolten in revealing the sites where electron transfer activities might take place, i.e. the HOMOs and LUMOs of the compounds. Fig. 11 and 12 respectively visualize HOMO and LUMO of the polyphenylene dendrimer.

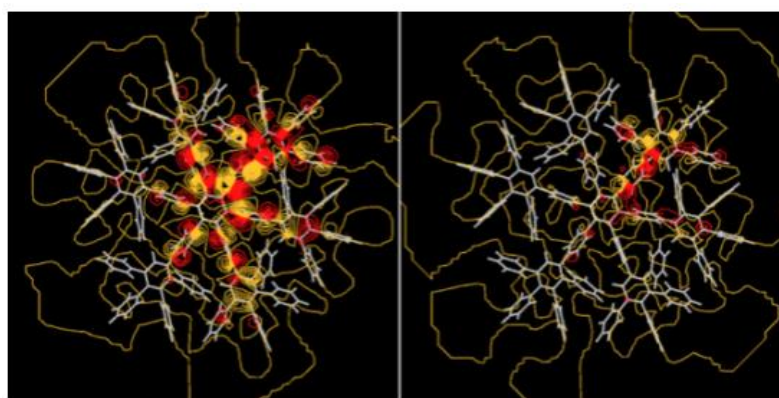


Figure 2. Molecular orbital HOMO and LUMO of the polyphenylene methyl dendrimer

In its most stable state, the polyphenylene methyl dendrimer has a frequency of electron transfer activities similar to that of chlorophyll. This was indicated by the existence of HOMO and LUMO in the polyphenylene dendrimer molecule. Therefore, it can be assumed that the polyphenylene dendrimer may act as chlorophyll substitute in artificial photosyntheses

Wavelength Analysis of Chlorophyll and Chlorophyll-like

The existence of HOMO and LUMO in chlorophyll molecules enables them to support electron transfer activities. Electron transfer is a key feature that singles out a potential chlorophyll substitute from the rest. This research has shown that the optimized structure of the polyphenylene methyl dendrimer possessed HOMO and LUMO. This indicates that the compound may support electron transfers as well as chlorophyll does

Table 1 Wavelength Comparison of Chlorophyll-like and Chlorophyll

Molecule	Wavelength	Spectrum
Polyphenylene Metyl Dendrimer	123-154 nm	UV
Ruthenium Complex	144-567 nm	UV-VIS
Chlorophyll	134-834 nm	UV-VIS

CONCLUSION

From the result data it can be concluded that the ruthenium complex is effective enough to perform in an artificial photosynthetic system as means of harvesting light and transferring electrons to break water molecules down into hydrogen and oxygen, yet not as effective as natural chlorophyll. According to the data obtained, the ruthenium complex is excited at UV (67 and 71 nm) and near infra red (943, 1188 and 1445 nm) rather than visible wavelength range. However, for the polyphenylene methyl dendrimer structure bears

8390.7265 joules of energy after being optimized using SVWN functional (a type of DFT functionals). From the molecular orbital analyses it was found that the dendrimer was capable of electron transfers as indicated by the existence of HOMO and LUMO and result comparisons with chlorophyll. Judging from the rate of technological advancement today, it is possible to develop this compound further in order to make it work better than or at least comparable to natural chlorophyll.

ACKNOWLEDGEMENT

We acknowledge funding from Research Loan Program 2023, Faculty of Mathematics and Natural Sciences, Mulawarman University, Ministry of Education and Culture, Republic of Indonesia.

REFERENCES

- [1] Curtis D. E. Lonie D. C. Vandermeersch T. Zurek E. Hanwell, M. D. and G. R. Hutchison. Avogadro: An advanced semantic chemical editor, visualization, and analysis platform. *Journal of Cheminformatics*, 4:17, 2012.
- [2] Frenking G. Kim K. S. Dykstra, C. E. and G. E. Scuseria. Theory and applications of computational chemistry. *Elsevier*, 2005.
- [3] Casewit C. J. Colwell K. S. Goddard W. A. Rappe, A. K. and W. M. Skiff. Uff, a full periodic table force field for molecular mechanics and molecular dynamics simulations. *Journal of the American Chemical Society*, 114 (25):10024–10035, 1992, doi:10.1021/ja00051a040.
- [4] Kongsted J. Crawford T.D. Pedersen, T.B. Gas phase optical rotation calculated from coupled cluster theory with zero-point vibrational corrections from density functional theory. *Chirality*, 21: E68–E75, 2009, doi:10.1002/chir.20778.
- [5] Stover J. Satman M. H. Williams J. Kiefte M. Kobly P. Son R. V. Burnier P. Bogusky D. Pfaff B., Darrington J. Gnu general public license. *Free Software Foundation, Inc. USA.*, 2007.
- [6] Baldrige K. K. Boatz J. A. Elbert S. T. Gordon M. S. Jensen J. H. Koseki S. Matsunaga N. Nguyen K. A. Su S. J. Windus T. L. Dupuis M. Montgomery J. A. Schmidt, M. W. General atomic and molecular electronic structure system. *J. Comput. Chem.*, 14:1347–1363, 1993.
- [7] D. C. Young. Computational chemistry: A practical guide for applying techniques to real-world problems. *John Wiley & Sons, Inc, New York.*, 2001 M. P. Brown and K. Austin, *The New Physique* (Publisher Name, Publisher City, 2005), pp. 25–30.

Laser Induced Fluorescence Imaging for Ripeness Detection of on-Tree-Oil Palm Fresh Fruit Bunches

Minarni Shiddiq⁽¹⁾, Alvika Juliardi Pratama⁽¹⁾, Zikri Radikel⁽¹⁾, Nadia Zakyyah Yasmin⁽¹⁾

⁽¹⁾*Department of Physics, Faculty of Mathematics and Natural Science, Universitas Riau
Jl. Soebrantas km 12.5, Simpang Baru, Pekanbaru, 28293, Indonesia*

**Corresponding author: minarni.shiddiq@lecturer.unri.ac.id*

Abstract. Laser-induced fluorescence imaging (LIFI) is one of the remote sensing methods for plant and fruit detection. Ripeness detection of on-tree-oil palm Fresh Fruit Bunches (FFBs) before harvesting is crucial for obtaining good quality FFBs. Traditional ripeness detection of FFBs uses surface color and several loosed fruitlets above ground. It is prone to subjectivity and difficulty for higher oil palm trees. This study proposed a LIFI system to detect the ripeness of on-tree oil palm FFBs. The LIFI system consisted of a monochrome CMOS camera with a 75 mm lens, two diode lasers, and optical filters. The laser wavelengths were 405 nm and 650 nm. Each on-tree FFB had six images, three parts (base, middle, top), illuminated by two lasers subsequently. The FFBs were then harvested for fruitlet firmness measurement. The correlations of fluorescence intensity, ripeness levels, and firmness were evaluated. Gray values of the image region of interest (ROI) represented the fluorescence intensity. Laser with 650 nm shows a higher correlation ranging from R2 of 0.52-0.87. The gray values decrease as the firmness level increases. The middle part of FFBs has the highest correlations. The results show the potential of the LIFI system for detecting ripeness of tree-oil palm FFBs.

Keywords: Laser induced fluorescence imaging, remote sensing, oil palm fresh fruit bunches, ripeness, firmness

INTRODUCTION

Many imaging methods have growing applications in agriculture, especially for fruit quality evaluations. Fruit qualities include external characteristics such as fruit surface color, shape, textures, and size, as well as internal such as chemical contents and physical attributes. Fruit ripeness or maturity is fruit quality represented by external or internal characteristics. The imaging techniques have two divisions: traditional computer vision and spectral imaging. Traditional computer vision is used widely for external characteristic evaluations since it has spatial resolution only. Spectral imaging techniques have spatial and spectral resolution, hence capable of evaluating physical-chemical attributes. Most spectral imaging methods used for fruit ripeness include hyperspectral imaging, multispectral imaging, fluorescence imaging, and laser light backscattering imaging. Other specific types of imaging techniques also used for fruit evaluations are Raman imaging, thermal imaging, infrared thermography, and microwave imaging [1]. The imaging techniques are non-destructive, fast, and reliable. They aim to substitute the manual and traditional methods to determine the fruit ripeness or maturity, such as using experienced graders. The manual and traditional methods are subjective, destructive, labor and time-inefficient, and tedious [2]. The imaging techniques accompanied by suitable machine learning have the potential to a rapid prediction of internal characteristics related to fruit ripeness.

The applications of imaging methods have been developed for the harvesting process. These applications have become part of the precision agriculture field of research. Precision agriculture uses recent technologies in computers and sensors, interdisciplinary fields to solve many problems in traditional agriculture with the growing world population and the need for new food production technologies. Fields of precision agriculture aim to reduce labor and operational costs through automation and real-time measurements. Precision Agriculture involves robotic harvesters and sensors for real-time measurements of soil nutrients, local weather, detections of plant diseases, yield estimations, and field management. These technologies are possible due to the fast development of computer vision, machine learning, and sensor systems [3]. Information on harvesting time is crucial because it can reduce pre- and post- harvest waste and increase yield. The imaging techniques play a role in predicting the harvesting time and crop, notably predicting fruit maturity or ripeness. Visible imaging or computer vision and spectral imaging are among the applicable imaging techniques used [4]. Another purpose of these efforts is to develop a robotic harvesting system for efficiency in time and labor cost [5].

Remote sensing techniques have found wide applications in evaluating and detecting intact fruit and vegetables during harvesting. Imaging methods such as computer vision, hyperspectral and multispectral imaging, fluorescence imaging, and light detection and ranging (LIDAR) have been explored for harvesting technologies, especially for maturity and harvesting period evaluation [6]. The methods rely on light sources as probes for the interaction of light and fruit materials. LIDAR is a type of remote sensing that uses electromagnetic waves in UV-Vis-IR regions, which is the opposite of RADAR, which uses a radio wave region. Traditional computer vision uses sunlight or other broadband light from halogen or light-emitting diode (LED) lamps. Hyperspectral and multispectral imaging usually use broadband light sources. However, a hyperspectral imaging system also included laser light for laser-induced fluorescence imaging in a comprehensive measurement of orange quality [7]. Imaging system using laser as a light source provides more advantages than those using other light sources because of its a long range, higher intensity, and good quality light beam.

Fluorescence imaging uses a fluorescence process that happens when light interacts with matter. The intensity and wavelength of the fluorescence image or signal can represent the information on the internal characteristics of matter. Chlorophylls and other pigments can be used as ripeness indicators [8]. Laser and light-emitting diode (LED) are often used as excitation light sources to induce fluorescence processes. Diode lasers and monochromatic LEDs are available commercially in optical power and wavelengths, which are compact, small, and cost-effective, hence can be integrated into portable devices. Methods of using fluorescence processes induced by laser or LED light are called laser-induced fluorescence (LIF). The detection method is called laser-induced chlorophyll fluorescence (LCIF) when chlorophyll of plant leaves or fruits is induced. When the detector is CCD or CMOS, it is called laser-induced fluorescence imaging (LIFI). LED is often used for short-distance detection because of its beam characteristics, short range, high divergence, and low optical brightness. Laser light has distinct properties from other light sources, such as low divergence, high brightness, and coherence. However, it has imminent dangers for human eyes and skin. Safety measures are crucial when dealing with lasers. The LIF or LIFI method has been used for many purposes, such as for tomato maturity stage discrimination [9], plant remote sensing monitoring [10], and estimating leaf chlorophyll content [11]. LIDAR instrument using a small industrial drone was also developed as the laser-induced fluorescence spectroscopy sensing of maize fields [12].

Indonesia is one of the largest exporter countries for crude palm oils (CPO). However Indonesian oil palm industries face many problems, such as low CPO qualities and plant diseases. Ripeness stages of oil palm fresh fruit bunches (FFBs) are the key factors to obtain good quality FFB and CPO. Good-quality FFBs are ripe FFBs, which have maximum total product content (oil) [13]. Such FFBs depend strongly on harvesting, sorting, and grading processes, which are carried out traditionally and manually. There have been many efforts to develop a cost-effective, reliable, and fast method of harvesting decision to obtain good quality oil palm fresh fruit bunch (FFB) before being sent to oil palm mills. Sorting and grading methods for oil palm FFBs at oil palm mills have been developed more intensively than harvesting methods. Traditional ripeness classification uses criteria on color change and the number of fruits loose from the bunch. The ripeness of oil palm FFBs is usually classified into three main stages, which are unripe, ripe, and overripe. Manual oil palm FFB sorting and grading need skillful graders who experience exhaustion, subjectivity, and tediousness for working for long periods.

There have been many efforts to evaluate or detect oil palm ripeness. The electronic, automatic, nondestructive, and reliable methods are preferable due to their effectiveness [14]. Computer vision is one of the nondestructive methods used for sorting and grading machine vision of oil palm FFBs, which uses a color camera and a computer accompanied by image processing software. Many image processing techniques for computer vision apply on color models to classify the FFB ripeness stages, such as red, green, blue (RGB), and hue, saturation, and intensity (HSI) color models [15]. A smartphone camera and image processing of color features were also used to segregate oil palm FFBs based on ripeness stages [16]. The other techniques were using fluorescence imaging [17], LIDAR [18], and triple flat-type air coil inductive sensor [19]. Destructive methods were performed for validation based on the physical, mechanical, and chemical properties of oil palm fruit, such as fruit firmness, oil content, flavonoid, and anthocyanin contents

Ripeness detection and classification techniques for on-tree oil palm FFBs before harvesting are less explored than sorting and grading methods for postharvest FFBs. Standard procedures for on-tree FFB harvesting are well established and used by oil palm farmers or industries. They are based on color change and the number of fruits loose from bunches and fell to the ground. The traditional methods always need experienced harvesters and have some flaws. Therefore, an electronic, cost-effective, and automatic system for ripeness detection is essential, especially for higher trees on which only small parts of an FFB appear. There have been some efforts made to simulate the maturity or ripeness detection of oil palm FFBs on-tree (remote sensing), such as using three lamps with different wavelength ranges and a color camera based on

RGB intensity [20] and laser-based computer vision with two laser wavelengths [21]. Both experiments still used harvested FFBS. Computer vision with eight color models was also used for oil palm FFB ripeness classification with a multi-class Support Vector Machine (SVM) [22]. Computer vision method using a color-based image processing algorithm has some drawbacks for on-tree (attached) fruits, especially for remote or high trees and sunlight environments.

This study aimed to develop a laser-induced fluorescence imaging (LIFI) system for ripeness detection of oil palm FFBS on trees. The built system was intended to be low-cost, automatic, miniature, and handy. It aimed for ripeness stage recognition of on-tree oil palm FFBS and to make the right decision for a harvester before harvesting oil palm FFBS. The system consists of a USB board CMOS, monochrome camera with a camera lens and a color filter, two diode lasers with 405 nm and 650 nm in wavelength, and a power supply for the diode lasers. The system was positioned in an acrylic box and mounted on a tripod for easy height and angle adjustments, brought to an oil palm plantation. Samples are on-tree oil palm FFBS, which were the candidates to be harvested by the plantation harvesters. In the image acquisition step, three fruits from each of the three parts of the candidate harvested FFB were illuminated by a focused laser beam sequentially for one second while its image was recorded by the camera. Then, the FFBS were harvested and brought to our laboratory for fruit firmness tests. The recorded images were processed using ImageJ to extract mean gray values [7, 8]. These gray values represented the fluorescence intensities. Correlation between the intensities and fruit firmness levels and the ratios of the fluorescence intensities at 405 nm and 650 nm were also analyzed.

MATERIALS AND METHODS

Sample Preparation

Samples of this study were on tree-oil palm fresh fruit bunches from the Nigresnce Tenera varieties, namely Tenera A and B, grown in most Indonesian oil palm plantations. Tenera A has larger fruits and less-density bunches. Tenera B has larger bunches and denser but smaller fruits. The on-tree oil palm FFBS were five trees with potential FFBS to harvest at the oil palm Laboratory plantation of the Faculty of Agriculture, Universitas Riau, on the harvest day. The ripeness stages of the FFBS before and after harvest were classified traditionally based on color change and the number of loose fruits by an experienced harvester. The ripeness stages of potential harvested FFBS (called fraction) on a tree consisted of F0 and F1 (unripe), F2 and F3 (ripe), and F4 (overripe). Each ripeness stage had 1 or 2 replicates depending on the harvest days. The heights of oil palm fresh fruit bunches observed were 2.5–3.5 m measured by a laser meter from the ground to a bunch middle part. Due to some circumstances in the plantations, FFB samples were obtained once a month, a total of 10 oil palm FFBS from 5 trees with five fractions. Each laser beam from the LIFI system was pointed consecutively to three fruits for three parts of an FFB bunch (top, middle, and bottom) located between two fronds of oil palm trees. Each oil palm FFB fruit with laser beam spot was recorded using the camera.

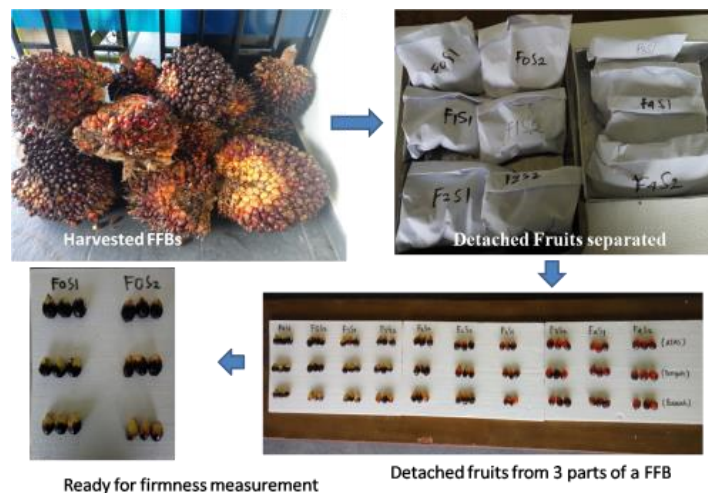


Figure 1. Harvested oil palm FFBS and detached fruit for firmness measurement

Figure 1 shows the process after image acquisition. After being illuminated by the laser beam and images recorded, the FFB samples were harvested on the same day by the plantation harvesters, brought to the Lab, cleaned from dirt, and weighted. The nine fruits from three parts of a bunch that have been illuminated and labeled were detached, collected, and relabeled according to their ripeness stages, as shown in Figure 1. The mass of each fruit was measured, followed by measuring their firmness levels by a manual penetrometer (Agriculture solution). We used the penetrometer tip of an 8 mm tip, intended for soft fruits (0.5 – 12 kg/cm²) in kg force/cm² unit. The firmness level for each part of a bunch is represented by the average firmness levels of three fruits. The firmness level of a detached oil palm FFB is the average of firmness values of the fruit's two sides.

Laser Induced Fluorescence Imaging System

This study aimed to use a laser-induced fluorescence imaging (LIFI) method to detect the ripeness of oil palm FFBs by analyzing the relations between the fluorescence intensities to the firmness levels of the FFB fruits. The firmness levels represented the ripeness levels of the FFB. Simple linear regression was used to analyze the relations.



Figure 2. LIFI System (a) system schematic, (b) system on tripod 2 m from a 3 m height of an oil palm tree

The laser induced fluorescence imaging system, as shown in Figure 2a, consists of two diode lasers with 405 nm (2) and 650 nm (3) of wavelength, respectively, a USB-board monochrome CMOS camera (5), two color filters and a 75 mm lens (6) [23]. Both lasers and the camera were positioned in a triangulation scheme to the FBB (1) and enclosed in an acrylic box (8), which can be mounted easily on a tripod. The lasers were supported by a power supply (4). During image acquisition, the system and tripod were brought to the plantations, positioned in front of a 3 m - potential oil palm tree, as shown in Figure 2b. For each possible oil palm tree, the system and tripod were located about 2.5 meters in front of the tree, 1 meter from the ground. All chosen candidate trees were about 3 m (2.5 m – 3.5 m) in height. All the height measurements used a digital laser level meter.

The main system components are the camera, color filters, and diode lasers, as shown in Figure 2b. The camera used was a small board CMOS camera (e-con system camera) which had a USB 3.0 data connection to a PC and image acquisition software. The camera has a board size of 40 mm × 44 mm, a maximum 1280 × 720 resolution at 60 fps, and an external trigger input. Smaller camera size allows cost-effective and smaller system, easily integrated to a microcontroller for stand-alone, automatic operation. The sensor in this camera has a 1/3 "form factor and a global shutter. The camera also has a CS-mount lens holder for easy exchanging suitable lenses. We used 75 mm, F/2.5 MP Machine Vision Lens, 2/3" Format. Two exchangeable color filters with different bandwidths were mounted in front of the camera. Both filters (Thorlabs.com) are two band-color filters, magenta (395 nm–440 nm, 645 nm–750 nm) for 650 nm laser and blue (390 nm–480 nm, 540 nm–750 nm) for 405 nm diode lasers. The filters used assess pigments and other chemical contents of the oil palm fruits, such as chlorophyll, anthocyanin, oil, and water [8, 17]. Both lasers used are Roithner diode lasers, a violet diode laser with a maximum optical power of 20 mW and a red diode laser with 24 m. Both lasers have TTL modulation capability.

Image acquisitions of oil palm FFB started after setting up the LIFI system on a tripod, and the candidate oil palm trees were marked. Due to the inflexibility of this system to electric sources, the measurements were held for oil palm trees close to electric sources, which limited the attached FFBs height varieties. We chose only around 3 m height of oil palm FFB trees with ages below ten years. Surrounding sunlight intensities were recorded every time an image was recorded. First, the 405 nm laser beam without an expander beam was aimed at one of the fruits at the middle part of a harvested potential FFB (parts seen between two fronds). An image of an FFB part illuminated by a laser beam was recorded using the camera, and saved in JPEG format. For each image taken, a spot image without a laser beam was also recorded. The procedure was repeated for other fruits in the middle part and further for fruits in other parts of an FFB. After all parts of potential FFBs were illuminated by a 405 nm laser beam, the procedures were then repeated for a 650 nm laser beam. These measurements were held in one-day period. Fruit firmness measurements then were continued after FFB harvesting.

Image Processing and Analysis

After all the images of fluorescence imaging were recorded and saved, average gray values were calculated using the open-source image processing software ImageJ 1.47v [24]. The method used the ROI of 50 pixels x 50 pixels for each image. Gray values were calculated from calibrated images (images with laser and without laser spot) to reduce the effect of sunlight intensity changes during measurement. The gray values represent the fluorescence intensities [8] of laser-induced fluorescence imaging applied to oil palm FFBs. The relationships of the fluorescence intensities and fruit firmness for each part of an oil palm FFBs and performance using a diode laser with different wavelengths were analyzed using linear regression.

RESULTS AND DISCUSSION

This study investigated the feasibility of using a laser-induced fluorescence imaging system to detect the ripeness of oil palm fresh fruit bunches (FFBs) on trees before harvesting. The system used two diode lasers with 405 nm and 650 nm wavelengths. The CMOS camera recorded the images of laser-illuminated FFBs. After harvesting, the ripeness of FFBs was classified traditionally based on color change and the number of fruits loose from bunches. Next, the ripeness stages were validated by fruit firmness levels of oil palm fruits, measured by a manual penetrometer. We presented more data results of Tenera A than Tenera B oil palm FFBs.

Oil Palm FFB fluorescence images

Figure 3 shows the images obtained by the LIFI system. The 650 nm laser beam illuminated a fruit in each part of a bunch, respectively. The images were recorded in JPEG format and processed using ImageJ software to extract gray values representing the fluorescence intensities. The calibrated images were the subtracting results of with and without laser beam (dark background) images. The calibrated image then obtained thresholding using the command adjust and threshold. Figure 3(top) shows images of each part of FFB lying between two fronds of the (top, middle, bottom) bunch. Figure 3 (below) shows the image processing results using ImageJ. The gray value for each part is the full-wave half maximum (FWHM) obtained from the program in arbitrary unit (a.u), which represents the fluorescence intensity of the oil palm FFB. The middle part has a higher gray value compared to the other parts. It is a potential representative part if one uses the method to detect FFB ripeness.

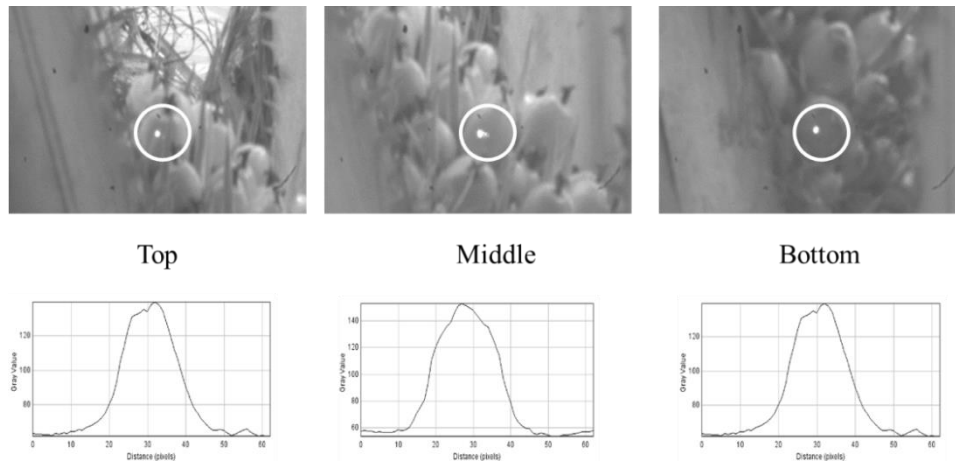


Figure 3. Images of ripe oil palm fruits illuminated by 650 nm with laser spots at three parts of an oil palm FFB tree.

Figure 3 shows the thresholding results of laser beam images on the FFB parts. The 650 nm laser beam (a) has larger diameter and intensities than the 405 nm laser beam (b). There is a slight difference in the thresholding diameter between FFB parts, could be due the angle and bunch parts of the fruit illuminated by the laser beam spot. A CMOS monochrome camera has some advantages than color camera for on-tree oil palm FFBs. It has broader wavelength range and higher quantum efficiency (QE). A monochrome camera has higher QE in near infrared region (NIR) and less affected by background light [17, 21]. Laser light gives more advantages for remote sensing of on-tree oil palm FFBs, such as long distance and more contrast on oil palm FFB images [20].

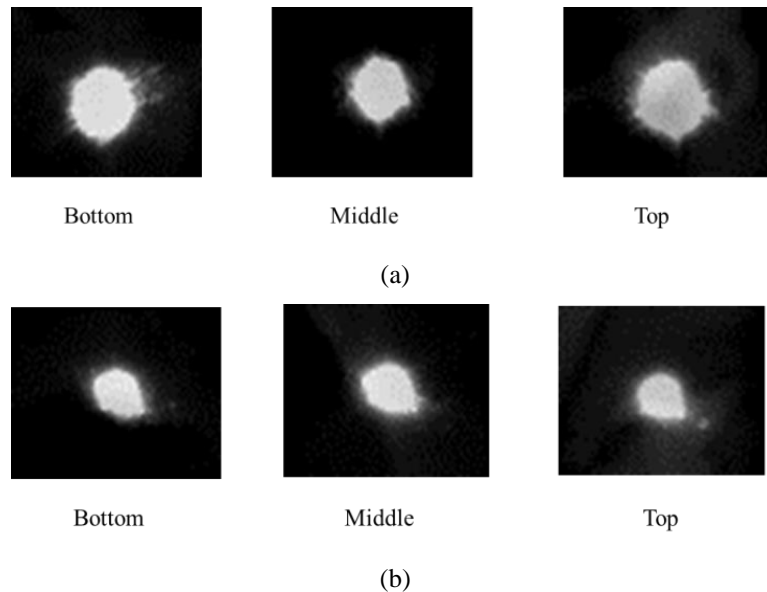


Figure 4. Images of 650 nm (a) and 405 nm (b) laser spots at three parts of an oil palm FFB tree

Oil Palm FFB Firmness to Ripeness Stages

Figure 4 shows the relationship between the fruit firmness (hardness) and ripeness stages (fractions), which are categorized traditionally right after harvesting. The measured fruit firmness was for fruits facing sunlight (fertile fruits), not those lying between fronds (infertile fruits). Figure 5 demonstrates that the fruit firmness decreases as the ripeness stages increase. The results were also revealed in another experiment using five steel balls with different diameters on CIRAD and COMPACT varieties of oil palm fruits. It was shown that the firmness difference between ripeness levels is related to the oil and moisture contents [25]. There is

also a slight difference in firmness for each part of a bunch, especially for unripe FFB (F0 and F1 fractions). The firmness values are the same for all bunch parts for ripe (F2) and over-ripe (F4) FFB.

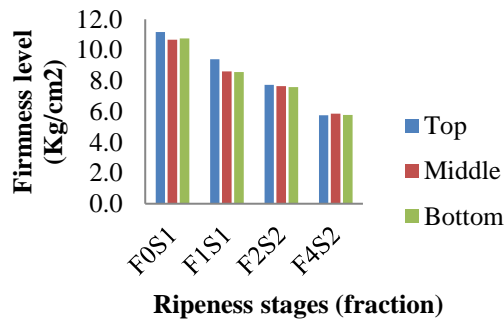


Figure 5. Relationship of fruit firmness values of each part of Tenera A oil palm FFBs to ripeness stages (fractions)

Relation of Fluorescence Intensity to Ripeness stages

Figure 5 displays the relation of fluorescence intensities versus the traditionally categorized ripeness of oil palm fruits detached after harvested, for 3 m tree height at different inducing laser beam. Laser with 650 nm shows more consistency that fluorescence intensity higher as the ripeness increases. This could relate to decreasing in chlorophyll and anthocyanin content, also increasing in oil content. Study of LIF application on papaya maturity has showed the chlorophyll fluorescence of papaya fruits having maxima; one in the red at 680–690 nm (F690), and the other in far-red region at 730–740 nm (F740) [9]. In this study, two filters were used, magenta (395 nm–440 nm, 645 nm–750 nm) for 650 nm laser, and blue (390 nm–480 nm, 540 nm–750 nm) for 405 nm diode lasers. Both filters are capable to record the reflected light for chlorophyll fluorescence.

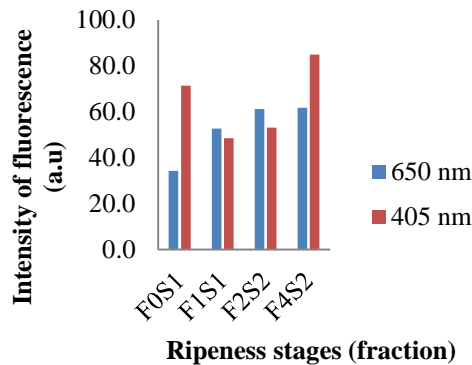


Figure 6. Fluorescence Intensities versus fruit firmness of oil palm FFBs using both diode lasers (Middle part)

Relation of Fluorescence Intensities to Fruit Firmness

The relation of fluorescence intensities to fruit ripeness was explored using two laser wavelengths and two filter types. The two filters are magenta (395 nm–440 nm, 645 nm–750 nm) for 650 nm (red) and blue (390 nm–480 nm, 540 nm–750 nm) for 405 nm (violet) diode lasers. Figure 7 shows the correlation of the fluorescence intensity to fruit firmness or hardness of Tenera A of oil palm FFBs using the red (a) and violet laser (b). It reveals that the fluorescence intensity for both lasers decreases as the firmness increases. It means unripe fruits with higher firmness have lower fluorescence intensities. Figure 7a has a higher correlation ($R^2 = 0.8576$) on the bunch's middle part, while Figure 7b has higher intensities ($R^2 = 0.6452$) on the bunch's top part. Laser-induced fluorescence (LIF) used to discriminate three different maturity stages of tomato showed that the peak intensities of LIF spectra decreased with the chlorophyll content depletion during ripening with

a regression coefficient of 0.85. Decreasing intensities as firmness increases are related to higher chlorophyll and anthocyanin contents [9] on unripe bunches. Higher firmness is also due to lower oil content [25].

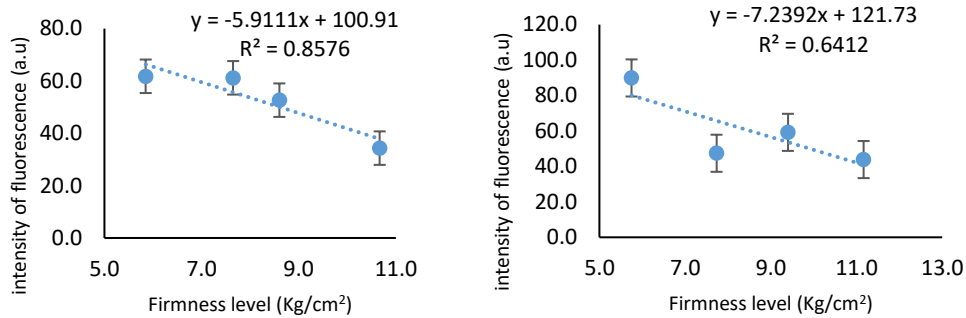


Figure 7. Fluorescence Intensities vs. fruit firmness values of oil palm FFBs using the 650 nm diode laser (middle part) (a) and 405 nm laser (top part) (b)

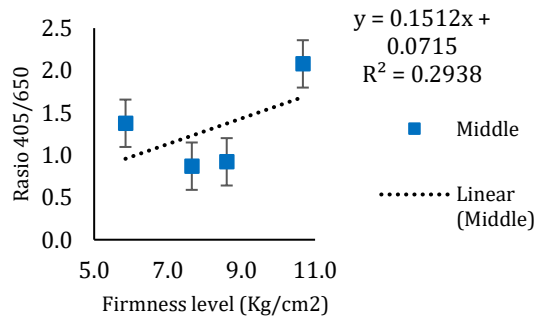


Figure 8. Ratio of fluorescence intensity measured using 405 nm laser and 650 nm laser for Tenera A oil palm FFB versus firmness levels (middle part)

Figure 8 shows the ratio of measured intensities by 405 nm laser and 650 nm for the middle part of a bunch versus firmness. The graph shows that the ratio values are higher for higher firmness levels. It means that the ratio decreases as the oil palm fruits ripen. This ratio was known as the Blue-to-Red Fluorescence (BRR_FRF) index. Experiment using fluorescence imaging obtained from blue-green (447 nm) and far-red (685 nm) emission signals showed a significant difference among the three categories of ripeness of oil palm FFB. In the same fashion, as shown in Figure 8, far-red fluorescence intensities are higher than the blue-green, resulting in a lower ratio of oil palm fruit ripening [26].

CONCLUSION

This study has successfully showed the possibility of using LIFI system for assessing oil palm FFBs on tree and detect the ripeness. Measurements of the FFB fruit firmness showed that the fruit firmness decreased as the ripeness stage increased. It is related to increasing oil content in the fruit mesocarp. A diode laser beam of 650 nm wavelength obtained a high correlation between fluorescence intensities and fruit ripeness levels of oil palm FFBs. It could be related to higher oil and water content as ripened fruits become softer. A diode laser of 405 nm showed a higher correlation of fluorescence intensities to pigment contents, chlorophyll, and anthocyanin. More images will be obtained for future work. Automatic image acquisition and processing will use a Python program. The system also needs artificial neural networks and user interfaces, hence can be used in the field. The height variety effects can be investigated if the portable instrument is used for higher trees. This study showed the potential of obtaining a low-cost, portable, automatic, Arduino, or Raspberry-based device used in the field for oil palm farmers or harvesters.

ACKNOWLEDGEMENT

This research was supported by Universitas Riau and funded partly by DRPM (Research Grant 478, 2017) and by BDPKKS (PRJ 33/DPKS/2018). The authors would like to thank Faculty of Agriculture, Universitas Riau for providing the FFB samples.

REFERENCES

- [1] S. Tian and H. Xu, *Food Engineering Reviews* 14, 380–407 (2022)
- [2] A. Hussain, H. Pu, and D. W. Sun, *Trends in Food Science and Technology* 72, 144–152 (2018).
- [3] A. Zujevs, V. Osadcuks, P. Ahrendt, *Procedia Computer Science*, 77, 227 – 233 (2015).
- [4] B. Li, J. Lecourt, and G. Bishop, *Plants* 7, 3 (2018).
- [5] E. Vrochidou, V. N. Tsakalidou, I. Kalathas, T. Gkrimpizis, T. Pachidis, and V. G. Kaburlasos, *Agriculture (Switzerland)*, 12, 1240 (2022).
- [6] H. Cakmak, “10- Assessment of fresh fruit and vegetable quality with non-destructive methods,” in *Food Quality and Shelf Life*, edited by C. M. Galanakis (Academic Press, 2019), pp. 303-331.
- [7] M. Liu, L. Zhang, and E. Guo, “Hyperspectral laser-induced fluorescence imaging for nondestructive assessing soluble solids content of orange,” In *Computer and Computing Technologies in Agriculture, Volume I. CCTA 2007, The International Federation for Information Processing 258*, edited by D. Li (Springer, Boston, MA, 2008), pp. 51-59.
- [8] E. N. Obledo-Vázquez and J. Cervantes-Martínez, *Appl. Opt.*, 56, 1753-1756 (2017).
- [9] L. M. Ali, R. H. El-Saeid, and M. Abdel-Harith, *J. Food Meas. Charact.*, 16, 3054–3061 (2022).
- [10] W. Wan and J. Su, *Meas. J. Int. Meas. Confed.*, 125, 564-571 (2018).
- [11] T. Hao, Y. Han, Z. Li, H. Yao, and H. Niu, *Appl. Opt.*, 59, 7734-7744 (2020).
- [12] V. N. Lednev et al., *Photonics* 9, 963 (2022).
- [13] I. John, A.-M. Magdalene, S. S. Syed Tarmizi, and J. T. Shirley, “A Model to Manage Crude Palm Oil Production System,” in *Engineering Application of Artificial Intelligence Conference 2018 (EAAIC 2018)*, MATEC Web Conf. 255, (EDP Sciences, 2019), p. 02001
- [14] S. L. Utom et al., *Int. J. Integr. Eng.*, 10, pp. 35-39 (2018).
- [15] M.K. Shabdin, A.R.M. Shariff, M. N. A. Johari, N. K. Saat, Z. Abbas, “A study on the oil palm fresh fruit bunch (FFB) ripeness detection by using Hue, Saturation and Intensity (HSI) approach,” in 8th IGRSM International Conference and Exhibition on Remote Sensing & GIS (IGRSM 2016), IOP Conf. Series: Earth and Environmental Science 37, (IOP Publishing, 2016), p.012039.
- [16] A. Ghazali, H. Selamat, Z. Omar, and R. Yusof, *Elektr. J. Electr. Eng.*, 18, 57–62, (2019).
- [17] H. Ishak, M. Shiddiq, R. H. Fitra, and N. Z. Yasmin, *J. Aceh Phys. Soc.*, 8, 84-89 (2019).
- [18] F. H. Hashim, Z. Mohd Zulkifli, T. Raj, and A. B. Huddin, *J. Kejuruter.*, 30, 93-101 (2018).
- [19] N. A. Aliteh, N. Mison, I. Aris, R. M. Sidek, K. Tashiro, and H. Wakiwaka, *Sensors (Switzerland)*, 18, p.2469 (2018).
- [20] D. Cherie, S. Herodian, U. Ahmad, T. Mandang, M. Makky, *International Journal on Advanced Science Engineering Information Technology* 5, 255–263 (2015).
- [21] M. Shiddiq, Fitmawati, R. Anjasmara, N. Sari, and Hefniati, “Ripeness detection simulation of oil palm fruit bunches using laser-based imaging System,” in *The 6th International Conference on Theoretical and Applied Physics (The 6th ICTAP) AIP Conf. Proc.*, 1801 (2017), pp. 050003-1–050003-7.
- [22] N. Sabri, Z. Ibrahim, and D. Isa, *Indones. J. Electr. Eng. Comput. Sci.*, 11, 549–557 (2018).
- [23] M. Minarni, Indonesia. Registered Patent No. P00201810737 (20 December 2018)
- [24] R. Fernández, C. Salinas, H. Montes, and J. Sarria, *Sensors* 14, 23885-23904 (2014).
- [25] T. Chucheep, N. Mahathaninwong, S. Limhengha, D. Petchui, and P. Templong, “Ripeness Inspection of Oil Palm Fruits by Applying Hardness Test Technique,” In *Proceedings of the 16th ASEAN Food Conference (16th AFC 2019)*, Outlook and Opportunities of Food Technology and Culinary for Tourism Industry 2020 (SCITEPRESS – Science and Technology Publications, Lda., 2022), pp. 279-283.
- [26] M. H. M. Hazir, A. R. M. Shariff, M. D. Amiruddin, A. R. Ramli, and M. Iqbal Saripan, *J. Food Eng.*, 113, 534-540 (2012).

Analysis of Soil Site Classes Based on Microtremor Measurements in the South Bengkulu, Indonesia

Arif Ismul Hadi^{*(1)}, M. Farid⁽¹⁾, Lindung Zalbuin Mase⁽²⁾, Refrizon⁽¹⁾, Shela Basaria Purba⁽¹⁾, and Darmawan Ikhlas Fadli⁽¹⁾

⁽¹⁾ *Department of Geophysics, Faculty of Mathematics and Natural Sciences, the University of Bengkulu, Jl. WR Supratman, Bengkulu, 38371, Indonesia*

⁽²⁾ *Department of Civil Engineering, Faculty of Engineering, the University of Bengkulu, Jl. WR Supratman, Bengkulu, 38371, Indonesia*

**Corresponding author: ismulhadi@unib.ac.id*

Abstract. This study aims to map earthquake-prone areas based on the soil site classes associated with a Vs30 value in the South Bengkulu, Indonesia. A microtremor measurement was performed. The microtremor data obtained is in the form of an H/V curve that is matched to the model curve, so that a shear wave velocity profile is obtained. From the shear wave velocity profile at a certain depth, the value of Vs30 is obtained at each study location. The results of the study show that the distribution of Vs30 values in the South Bengkulu region, Indonesia consists of two soil site classes, namely site class C and site class B. Site class C (SC) consists of very dense soil and soft rock, while site class B (SB) consists of rock. The SC soil site class is distributed in a southwest-northeast direction and the SB soil site class is more dominant in a southeast-northwest direction. Based on the class of the soil site, this area is considered a relatively safe area from the threat of an earthquake. However, it should be noted for the lower shear wave velocity values at several study locations.

Keywords: site class, microtremor, shear wave, Vs30, earthquake.

INTRODUCTION

South Bengkulu is part of the Bengkulu Province on the Indonesian island of Sumatra. Geographically, the South Bengkulu region is in the subduction zone region and is traversed by the Manna segment Sumatran fault [1–5]. This condition results in the South Bengkulu region being vulnerable to earthquakes, both originating from the sea and from the mainland. Until now, earthquakes cannot be predicted accurately [3,6]. Therefore, efforts that can be made to anticipate it is through disaster mitigation. One of the steps that can be taken to mitigate the earthquake is to know the soil site classes at the study site. The soil site class describes the general condition of the soil and is used to define the amplification factor at the site [7,8]. This soil site class is determined based on the average shear wave velocity to a depth of 30 m (Vs30) [9]. The amplification factor caused by earthquakes can affect up to a depth of 30 m [10]. For this reason, it is necessary to know the Vs30 value for determining the soil site classes at the study site as one of the earthquake disaster mitigation efforts.

Determination of the Vs30 value can be carried out using several geophysical methods, for example the Multichannel Analysis of Surface Wave (MASW) method and microtremor measurements. Determination of the value of Vs30 through the MASW technique has been carried out by several researchers, such as [11–15] and others, while microtremor measurements were carried out by [16,17]. In this study, we investigated Vs30 using microtremor measurements. Microtremor measurement is a passive seismic geophysical method. Seismic sources come from ambient vibrations in the form of vibrations from nature, for example oceans and from humans, for example traffic, industry, and human activities [17]. This study aims to map earthquake-prone areas based on soil site classes related to the Vs30 value in the South Bengkulu region, Indonesia. The results of this research can be used as guidelines for disaster mitigation in the study area.

MATERIALS AND METHODS

This study was conducted in South Bengkulu, Bengkulu Province, Indonesia. The amount of study data is 104 measurement points. The distribution of microtremor measurement data points is shown in Figure 1. The tool used for microtremor data acquisition is a PASI Gemini-2 Seismometer (triaxial geophone) to record seismic waves which are directly stored on a laptop. Each microtremor measurement point is recorded using the Global Positioning System (GPS).

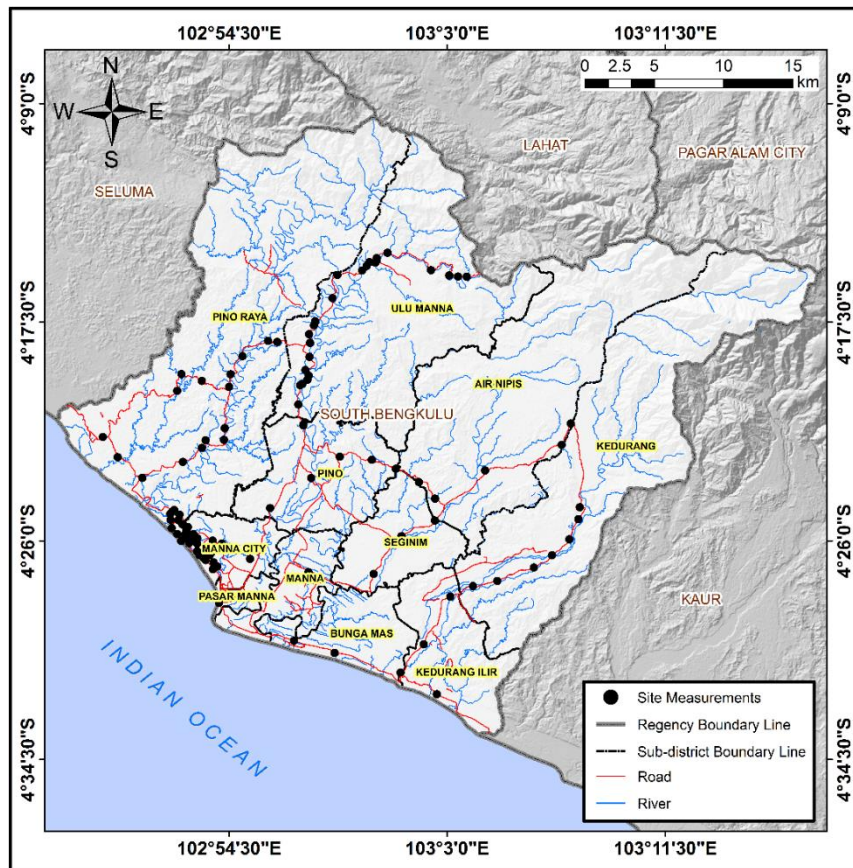


Figure 1. Distribution of microtremor measurement data points in the study area.

This microtremor data collection refers to the Site Effects Assessment using Ambient Excitation (SESAME) standard. The long duration of data collection will provide the possibility of low predominant frequency data output [18]. For this reason, the duration of microtremor data collection was carried out for 30 minutes at each study location. After the microtremor data is obtained in the field, data processing is then carried out to obtain the predominant frequency value of the soil and its amplification factor. Then all the results of this data processing are tested for reliability according to the requirements of SESAME. If the reliability test results are met, then the H/V curve is matched with the model curve to obtain a shear wave velocity profile at a certain depth. This result will be good, if the misfit curve between the field data and model data is close to one another. The shear wave velocity profile at a certain depth is then determined by the value of Vs30 from the relationship [9,19]:

$$\bar{V}_s = \frac{\sum_{i=1}^n d_i}{\sum_{i=1}^n \frac{d_i}{V_{si}}} \quad (1)$$

where d_i is the thickness of the i -th layer (m) and V_{si} is the shear wave velocity in the i -th layer (m/s). After obtaining the Vs30 value at each study location, then data analysis and interpretation were carried out. Analysis and interpretation of data is based on the classification of soil site classes from the Indonesian National Standard [19] as shown in Table 1. The denser a soil, the greater its Vs30 value and vice versa.

Table 1 Classification of soil sites based on SNI, 2019 [19]

Site Classes	Soil Types	The average shear wave velocity, \bar{V}_s (m/s)
SA	Hard rock	$\bar{V}_s \geq 1500$
SB	Rock	$750 < \bar{V}_s \leq 1500$
SC	Very dense soil/soft rock	$350 < \bar{V}_s \leq 750$
SD	Stiff soil	$175 < \bar{V}_s \leq 350$
SE	Soft soil	$\bar{V}_s < 175$

RESULTS AND DISCUSSION

The microtremor data obtained in the field consists of seismic signals on a vertical component (Up-Down/U-D) and a horizontal component (North-South/N-S and East-West/E-W). The microtremor data is then selected for signals that do not come from noise. It is characterized by a small amplitude rather than a transient vibration. Sampling of the window length is carried out every 15 s, so that more seismic signal measurements are obtained. An example of microtremor data on the U-D, N-S, and E-W components from the field is shown in Figure 2.

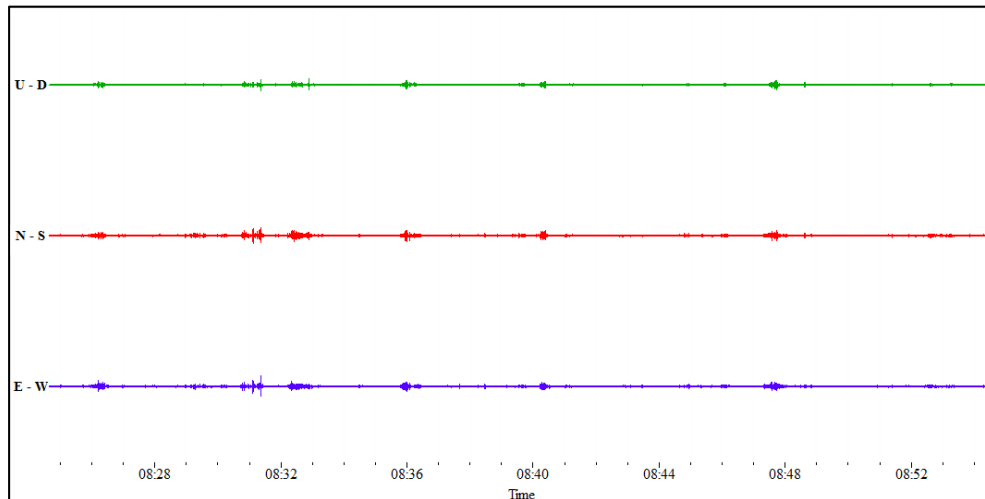


Figure 2. Examples of microtremor data recording signals on vertical components (U-D) and horizontal components (N-S and E-W)

In Figure 2 the results of the microtremor data on these measurements are good and there is very little noise, so it is very easy to analyze. After selecting a good signal, it is then processed further using Geopsy software. The results of this Geopsy processing can obtain the value of the predominant frequency of the soil and its amplification factor. An example of Geopsy processing results is shown in Figure 3.

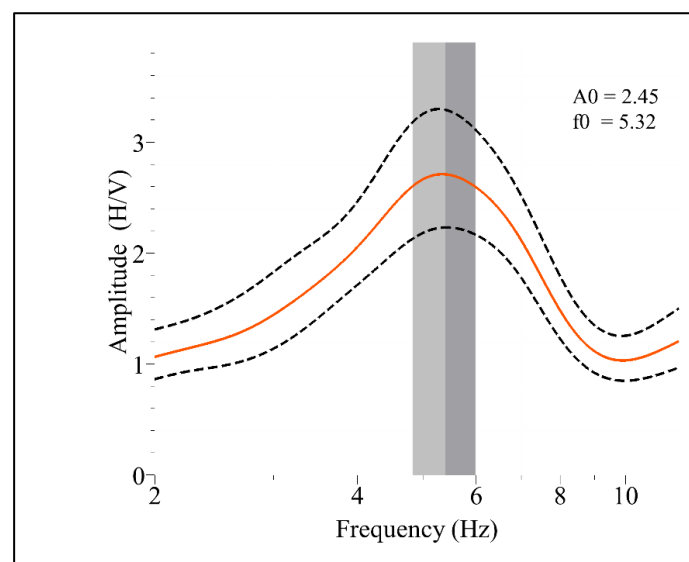


Figure 3. The results of data processing using Geopsy with a predominant frequency value (f_0) = 5.32 Hz and amplification factor (A_0) = 2.45.

Based on the reliability analysis, all the microtremor field data met the reliability criteria of [18], so these results can be continued to obtain shear wave velocity values (V_s) per rock layer by matching the H/V curve from field data with the calculated model curve. An example of a comparison of the H/V curve from field data with the model curve from the results of calculations through the inversion process is shown in Figure 4.

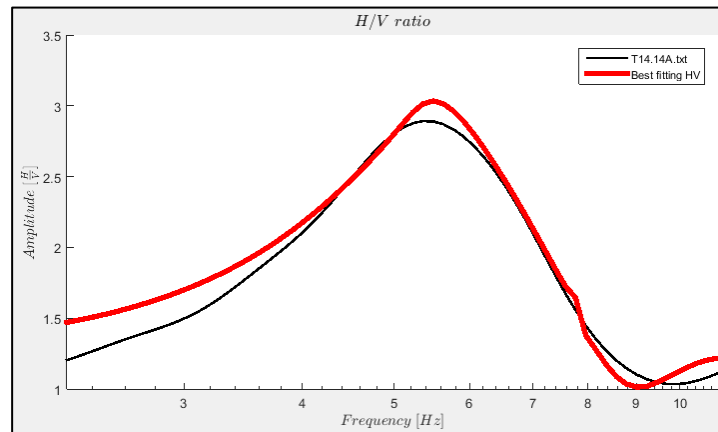


Figure 4. Comparison of the field data H/V curve and the model curve from the calculation results. The black lines are field data and the red lines are model data.

In Figure 4 the resulting misfit is quite tangent between the H/V curve from the field data and the model curve from the calculation data. The results of this processing indicate that the data obtained is good. The closer the H/V curve from the field data to the model curve from the calculation data, the better the data will be and the farther away from it, the worse the data will be. From the results of V_s per rock layer for each research location, then the value of V_{s30} is obtained through Equation (1). An example of the V_s value per rock layer and the V_{s30} value can be seen in Figure 5.

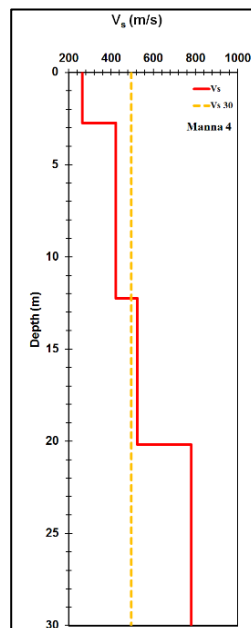


Figure 5. Example of value and rock layers V_{s30} value at Manna 4 location. The red line is the value of V_s per rock layer and the dotted orange color line is the value of V_{s30}

In Figure 5, the values for the rock layers in the first, second, third, and fourth layers are: 263 m/s, 424 m/s, 526 m/s, and 781 m/s, while the value of V_{s30} is 496 m/s. The distribution of V_{s30} values in the study area is shown in Figure 6.

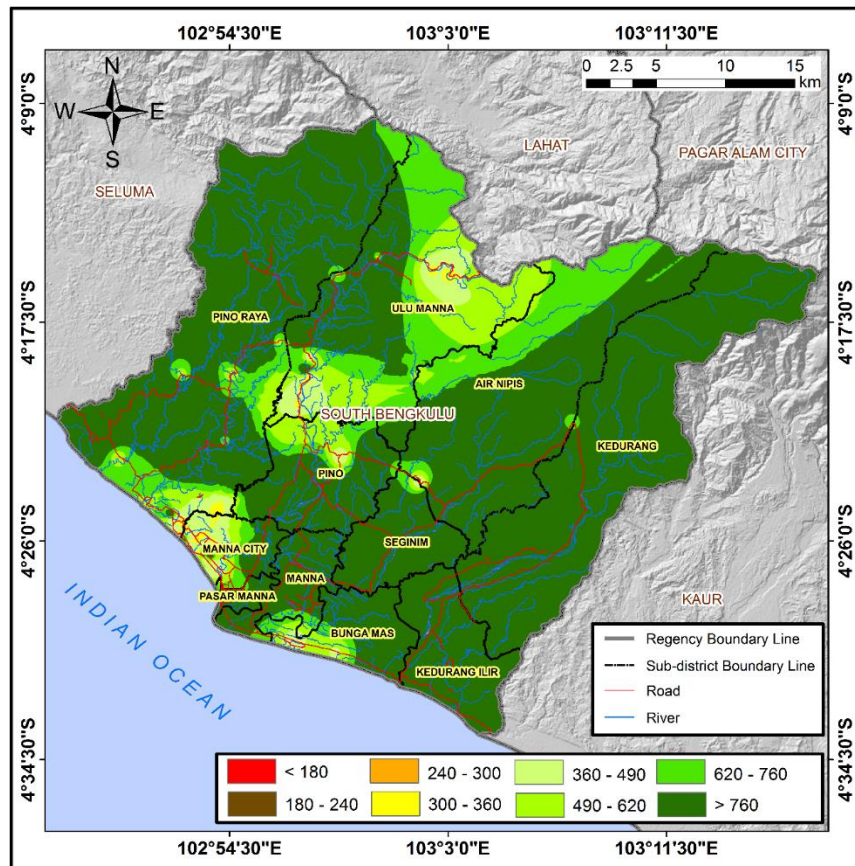


Figure 6. The distribution of V_{s30} values in the South Bengkulu region uses microtremor measurements

The V_{s30} class range is divided according to the class division of [20]. The lowest V_{s30} value is 268 m/s and the highest is 1457 m/s. The low shear wave velocity values are distributed in the southwest and northeast, while the high V_{s30} values are mostly distributed in the southeast and northwest. Low V_{s30} values are distributed in the southwest and northeast areas, while high V_{s30} values are mostly distributed in the southeast and northwestern regions. Furthermore, from the V_{s30} value at each study location, the distribution of soil site classes was mapped according to SNI classification [19]. Soil site class A, is hard rock (SA) with shear wave velocity > 1500 m/s, Soil site class B, is rock (SB) with shear wave velocity between 750 m/s to 1500 m/s, soil site class C, is very dense soil or soft rock (SC) with shear wave velocities between 350 m/s to 750 m/s, soil site class D, is moderate soil (SD) with shear wave velocities between 175 m/s to 350 m/s, and soil site class E, is soft soil (SE) with a shear wave velocity < 175 m/s [19]. The distribution of soil site classes in the South Bengkulu region can be seen in Figure 7.

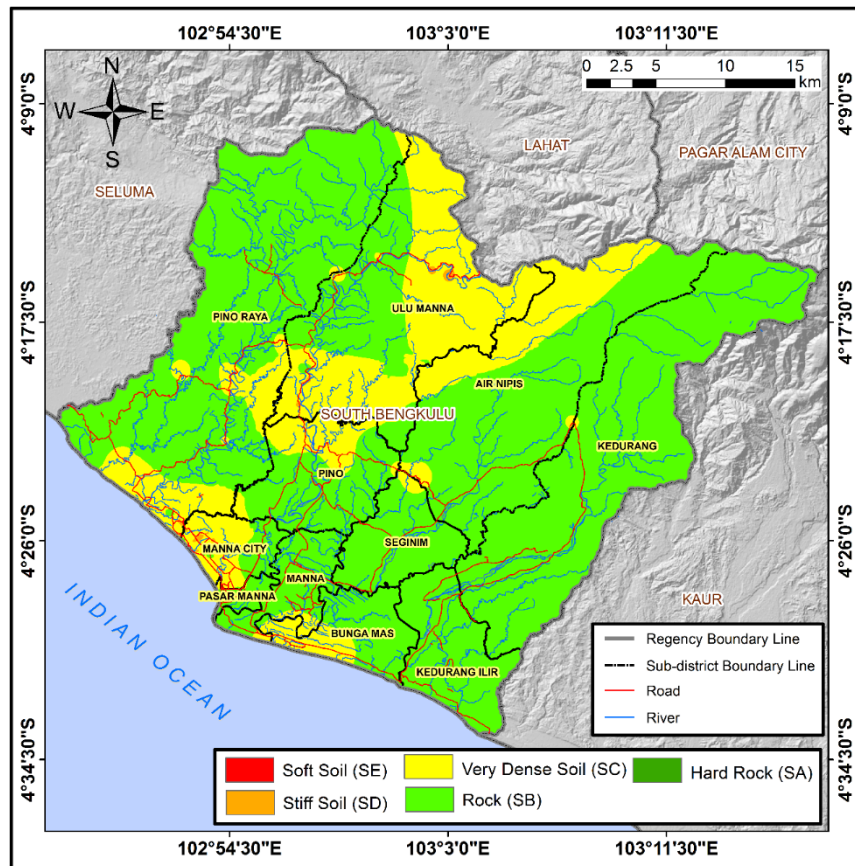


Figure 7. The distribution of soil site classes based on microtremor measurements in the South Bengkulu region.

In Figure 7 it can be shown that in the South Bengkulu region, Bengkulu Province, Indonesia is divided into two soil site classes, namely very dense soil (SC) and rock (SB) sites. The distribution of soil site classes in the study area is in accordance with the distribution of V_{s30} values, namely the SC soil site class is distributed in the southwest-northeast direction and the SB soil site class is more dominant in the southeast-northwest direction. The shear wave velocity value is related to the distribution of damage in a region due to an earthquake. The lower the shear wave velocity in a region, the greater the distribution of damage due to earthquakes and vice versa [11]. The elastic properties of near-surface rocks and their influence on seismic wave propagation are very important for earthquake studies, civil engineering, seismic risk, seismic hazard zonation, earth science and environmental studies [11,21]. The amplification of seismic waves will be even greater due to the presence of soil deposits near the surface [22]. Soft soil will provide a higher seismic vibration response compared to harder soil when an earthquake occurs [12,16]. Based on the 2000 earthquake in Bengkulu with a magnitude of 7.3 from Meteorology, Climatology, and Geophysical Agency (BMKG) data, the South Bengkulu region had a lower level of damage compared to other areas which were located farther from the epicenter of the earthquake. This can be explained that the South Bengkulu region consists of very dense soil and rock associated with a low level of amplification. This region is relatively safer from the threat of earthquakes related to the level of distribution of damage to a building.

CONCLUSION

Soil classification based on the V_{s30} value in the South Bengkulu region, Bengkulu Province, Indonesia consists of a very dense soil (SC) site class and a rock soil (SB) site class. The SC soil site class is generally distributed in a southwest-northeast direction, while the SB soil site class is generally distributed in a southeast-northwest direction. Based on the soil site classes, these regions are relatively safe from earthquake threats. However, it should be noted for the lower shear wave velocity values at several study locations. The results of this study can also be used as a reference for earthquake disaster mitigation at the study location.

ACKNOWLEDGEMENT

This research was supported by funding sources from "Penelitian Unggulan Universitas Bengkulu 2023" with the contract number: 2014/UN30.15/PP/2023. We thank the Institute for Research and Community Service at the University of Bengkulu for its support.

REFERENCES

- [1] M.D. Petersen, S. Harmsen, C. Mueller, K. Haller, J. Dewey, N. Luco, A. Crone, D. Lidke, and K. Rukstales, *Documentation for the Southeast Asia Seismic Hazard Maps* (U.S. Geological Survey, Virginia, 2007).
- [2] K. Megawati, T-C. Pan, and K. Koketsu, *Soil Dynamics and Earthquake Engineering* **25(1)**, 11–25 (2005).
- [3] PUSGEN, *Peta Sumber dan Bahaya Gempa Indonesia Tahun 2017* (Pusat Studi Gempa Nasional, Pusat Litbang Perumahan dan Permukiman, Badan Penelitian dan Pengembangan, Kementerian Pekerjaan Umum dan Perumahan Rakyat, Bandung, 2017).
- [4] A.I. Hadi, M. Farid., Refrizon, B. Harlianto, N. Hidayat, M. Kribudianto, *Jurnal Fisika Flux* **18(2)**, 105–118 (2021)
- [5] I.A. Awaliyah, A.I. Hadi, M. Farid, D.I. Fadli, A.J. Akbar, Refrizon, *Jurnal Penelitian Pendidikan IPA (JPPIPA)* **9(2)**, 592–599 (2023).
- [6] M. Irsyam, W. Sengara, F Aldiamar, S. Widiyantoro, W. Triyoso, D.H. Natawidjaja, E. Kertapati, I. Meilano, Suhardjono, M. Asrurifak, and M. Ridwan, *Ringkasan Hasil Studi Tim Revisi Peta Gempa Indonesia 2010* (Kementerian Pekerjaan Umum, Bandung).
- [7] S. Yaghmaei-Sabegh, and B. Hassani, *Soil Dynamics and Earthquake Engineering* **128**, 1–6 (2020).
- [8] T.I. Allen and D.J. Wald, *Topographic Slope as a Proxy for Seismic Site-Conditions (V_s^{30}) and Amplification Around the Globe* (U.S. Geological Survey, Alaska, 2007), pp. 69.
- [9] BSSC, NEHRP Recommended Provisions for Seismic Regulations for New Buildings and Other Structures (FEMA 450), Part 1: Provisions, Building Seismic Safety Council National Institute of Building Sciences (BSSCNIBS, Washington, D.C., 2004).
- [10] W. Wangsadinata, *Perencanaan Bangunan Tahan Gempa Berdasarkan SNI 1726 – 2002* (Shortcourse HAKI, Jakarta, 2006).
- [11] A.I. Kanli, P. Tildy, Z. Pronay, A. Pinar, and L. Hermann, *Geophys. J. Int.* **165**, 223–235 (2006).
- [12] M. Muzli, R.P., Mahesworo, R. Madijono, Siswoyo, S. Pramono, K.R. Dewi, Budiarta, O. Sativa, B. Sulisty, R. Swastikarani, N. Oktavia, Moehajirin, N. Efendi, T.A. Wijaya, B. Subadyo, Mujianto, Suwanto, and S. Pramono, *Jurnal Meteorologi dan Geofisika* **17(1)**, 25–32 (2016).
- [13] Y. Alan, M. Antony, B. John, *Soil Dynamics and Earthquake Engineering* **127**, 1–13 (2019).
- [14] D. Putri, Refrizon, A.I. Hadi, Penentuan Nilai V_s Per Kedalaman di Pesisir Pantai Panjang Kota Bengkulu Menggunakan Metode Multichannel Analisis Of Surface Wave (MASW), *Prosiding Seminar Nasional Fisika 7.0*, 270-276 (2021).
- [15] M. Prasisila, Refrizon, A.I. Hadi, Klasifikasi Kelas Situs Tanah Dengan Nilai V_s30 Di Kecamatan Muara Bangkahulu Kota Bengkulu Menggunakan Metode Multichannel Analysis Of Surface Wave (MASW), *Prosiding Seminar Nasional Fisika 7.0*, 277-282 (2021).
- [16] Nurrahmi, R. Efendi, and Sandra, *Gravitasi* **14(1)**, 7–12 (2015).
- [17] Y. Nakamura, Clear Identification of Fundamental Idea of Nakamura's Technique and Its Application, *The 12th World Conference on Earthquake Engineering*, January 30–February 4, 2000, Auckland, New Zealand.
- [18] SESAME, Guidelines for the Implementation of the H/V Spectral Ratio Tehnique on Ambient Vibrations: Measurements, Processing and Interpretation, European Commission – Research General Directorate, Project No. EVG1-CT-2000-00026 (SESAME, Brussels, 2004), pp. 8–31.
- [19] SNI, Tata Cara Perencanaan Ketahanan Gempa untuk Struktur Bangunan Gedung dan Non Gedung, Badan Standardisasi Nasional (SNI 1726:2019) sebagai Revisi dari SNI 1726:2012 (Badan Standardisasi Nasional, Jakarta, 2019).
- [20] T.I. Allen and D.J. Wald, *Bulletin of the Seismological Society of America* **99(2A)**, 935–943 (2007b).
- [21] J. Martin and J.G. Diehl, Practical Experience Using a Simplified Procedure to Measure Average Shear-Wave Velocity to a Depth of 30 Meters (V_s^{30}), *13th World Conference on Earthquake Engineering*, August 1-6, 2004 Vancouver, B.C., Canada.
- [22] V.W. Lee and M.D. Trifunac, *Soil Dynamics and Earthquake Engineering* **30(10)**, 1250-1258 (2010).

Investigation of Landslide in the Ulu Manna, South of Bengkulu District Based on Horizontal to Vertical Spectral Ratio Inversion

Muhammad Farid^{*(1)}, Arif Ismul Hadi⁽¹⁾, Fadillah Humairah⁽¹⁾

⁽¹⁾*Geophysics Department, Mathematics and Natural Science Faculty - Bengkulu University - Indonesia*

**Corresponding author: mfarid@unib.ac.id*

Abstract. The value of V_s30 as an important parameter in predicting the movement of seismic waves. Data V_s30 can be obtained from microtremor data and ultrasound data. This research aims to identify avalanche zones based on HVSR inversion in Ulu Manna District, South Bengkulu in mitigating landslide disasters in the area. The amount of data used was 30 microtremor data measurement points. Microtremor data was processed using Geopsy software which was analyzed using the HVSR system, HVSR inversion was carried out using Hv-Inv software and value of V_s30 compared to a value of V_s30 from USGS data to obtain accurate results. The results of microtremor data research showed that the value of V_s30 in Ulu Manna District ranged from 243,96 m/s - 766,84 m/s and the soil site class was dominated by class D with medium soil texture which was classified as a high seismic vulnerability zone. The ratio of V_s30 from the results of microtremor data and USGS topographic models varies around 0.56 - 1.76. This shows that the results of data V_s30 microtremors are different, but not much different in value from data V_s30 USGS. Based on the results of microtremor data research conducted in the research area coupled with the results of rainfall data (2013 - 2022) in Ulu Manna District which is very high with a total of 3854 – 5513 mm per year, Ulu Manna District is classified as an area prone to potential landslides.

Keywords: V_s30 , Microtremor, USGS, Seismic Vulnerability Zone, Avalanche.

INTRODUCTION

South Bengkulu Regency is one of the regencies in Bengkulu Province which has uneven geomorphological conditions [1], and is an earthquake-prone area [2]. In direct observation, geomorphological conditions in South Bengkulu district are very common with cliffs and ravines. Cliffs and ravines in earthquake-prone areas will be very easy to move, which on a large scale will result in landslides [3]. Several researchers have studied the phenomenon of landslides throughout the world, both caused by the influence of seasons and by earthquake shocks [4]. In general, landslides can be triggered by earthquakes and heavy rainfall [5], [6]. In the other studies, it was stated that earthquake shocks can trigger tsunamis, liquefaction and landslides [7]. In addition to earthquakes and rain on a large scale, areas with cliffs have the potential for landslides [8]. The potential for landslides is greater if the slope angle of the cliff is greater with respect to the horizontal plane [9]. Cliffs and ravines both have the potential to experience landslides, because the cliffs and ravines form an angle that allows soil material to move. [10] stated that landslide-prone areas occur on cliffs and ravines with a slope angle between 25°- 45°. In research on the physics of the coast, it was argued that areas prone to landslides also occur on coastal cliffs [11].

Landslides are a complex phenomenon and require a variety of surveying tools (e.g. drilling, inclinometers, INSAR, geophysics, and global positioning systems). Non-destructive classical geophysical exploration methods (e.g. seismic refraction, electrical resistivity tomography) are widely applied in slope stability studies [12]. However, this method is expensive, time consuming, and challenging to implement, especially in difficult terrain. This limitation can be overcome by applying the Horizontal-to-Vertical Spectral Ratio (HVSR) Method, which is easier, cheaper, and quicker to implement than the methods above. This technique has previously been used worldwide to analyze unsteady regions under spectral response analysis [13]. In the microtremor data survey method known as Horizontal to Vertical Spectral Ratio (HVSR) analysis, it is possible to identify the geological characteristics of an area [14]. HVSR data analysis can find shear wave velocity values up to a depth of 30 meters (V_s30).

Several areas in Indonesia have experienced landslides as a result of small and large scales earthquakes, due to the movement of tectonic plates [15]. Pressure on mineral elements and weak areas in rock and soil caused by earthquake vibrations causes slopes to become unbalanced and can trigger landslides [16].

Other alternative is to obtain shear wave velocity value data by applying a model from USGS data based on topographic modeling [17]. USGS data can be accessed and can be compared with values from microtremor data. Solid, hard and massive rocks, for example igneous rocks, have a low landslide susceptibility factor. Rocks that are not dense, soft and easily eroded, for example clay, shale and siltstone, have a high level of landslide susceptibility factor [18]. The Ulu Manna causeway is a road that connects

Bengkulu Province with South Sumatra Province. If a landslide occurs at one point on this road section, all transportation of goods and people will be hampered, both from Bengkulu to South Sumatra, or vice versa. The possible losses that occur are not only material and human souls, but also time [19]. If there is no alternative way, then the position of this road is very important. To avoid landslides that occur suddenly, it is necessary to know and map locations that are geologically very prone to landslides that can occur suddenly. Landslides that can occur suddenly are due to the condition of the slip surface which has a flat morphology [20]. The same thing was stated by [21] that the stability of ground motion is strongly influenced by the surface area of the slip plane.

In addition to morphological conditions, the angle of inclination also determines the occurrence of sudden landslides [22]. Landslides can also occur suddenly due to local geological conditions and rainfall that exceeds the average capacity that has ever occurred in the region [23]. A map of locations that are very prone to landslides on the Ulu Manna causeway space, has just been made based on shear wave velocity (V_s).

The landslide hazard map based on the speed of the new shear wave is a preliminary illustration for its accuracy. On that basis, it is very important to do mapping involving other variables that will strengthen the accuracy of the map. Other variables to be studied are the morphological conditions of the slip surface based on microtremor data, the slope angle of the slip plane, local geological conditions, and the cover biota of the cliff slope.

MATERIALS AND METHODS

Materials

This research does not use special materials, because the data will be obtained directly from the tool. The data acquired is microtremor data in the form of ground vibration data. Ground vibration data is commonly referred to as ambient noise, which is very small harmonic vibrations that occur continuously, are trapped in the surface sediment layer, reflected by the presence of layer boundaries with a fixed frequency, caused by micro-vibrations under the soil surface and other activities. This data is then processed using Geopsy software with the Horizontal to Vertical Spectral Ratio (HVSr) method. From the data processing, a spectrum is obtained that describes the dominant frequency value (f_0) in the horizontal direction and the H/V value as the amplification factor (A_0), as shown in Figure 1.

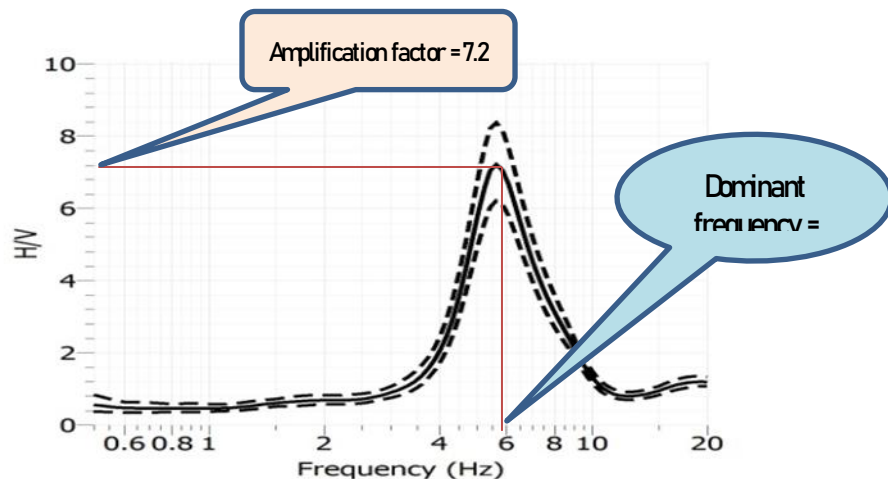


Figure 1. Example of H/V spectrum and dominant frequency

This study uses the main tool in the form of a 3-component short period seismometer, and is supported by other tools used in data acquisition as shown in Table 1.

Table 1. Data acquisition tools and their functions

No.	Tools and materials	Function
1.	<i>Seismometer short periode 3 component</i>	Records ground vibration data
2.	Laptop	Processing data

No.	Tools and materials	Function
3.	Compass	Signpost
4.	Global Positioning System (GPS)	Coordinate pointer
5.	Software Geopsy	Process data into information
6.	Microsoft Office Excel	Do the calculations
7.	Software ArcGis (ArcMap) 10.3	Create a map
8.	Software HV-Inv	Do the inversion

Data acquisition is designed according to the conditions in the field. The amount of data acquired was 30 points with a duration of 30 minutes for each point. Determination of the location of data acquisition is carried out as shown in Figure. 2 following:

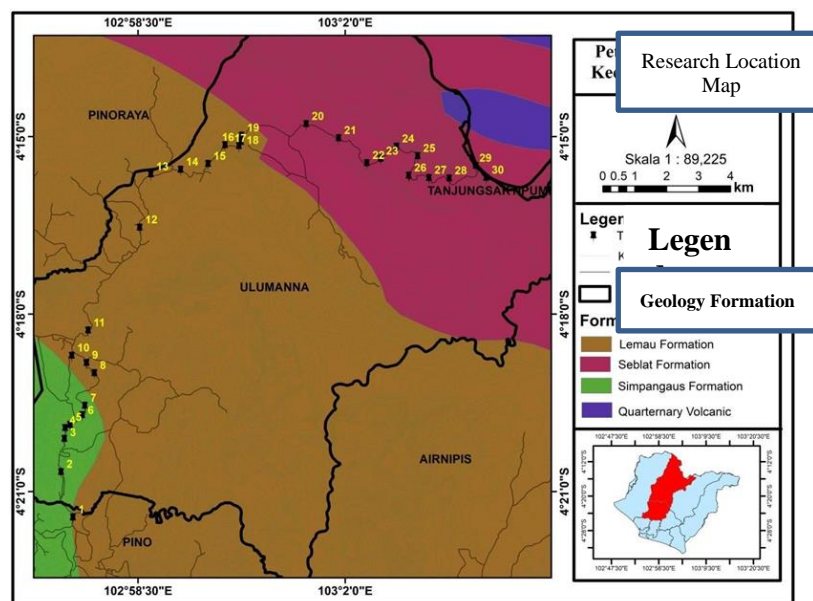


Figure 2. Map of the distribution of V_{s30} values in Ulu Manna District, South Bengkulu.

Methods

The method used in this research is the passive seismic method, which is a method that uses seismic waves generated by nature with low frequencies that can be used for monitoring earthquake predictions, and for monitoring geothermal systems.

The sequence of this research is shown in the flow chart as shown in Figure 3.

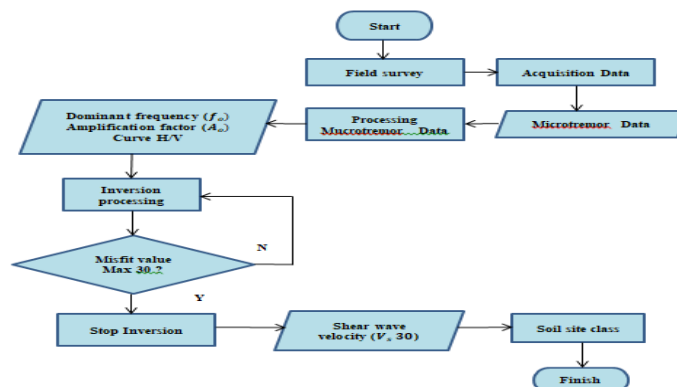


Figure 3. Flow Charta Acquisition and processing microtremor data to obtained V_{s30} and Soil ste class

The shear wave velocity value (V_{s30}) is obtained from the results of the Horizontal to Vertical Spectral Ratio inversion using the HV-Inv software. The value of V_{s30} from the results of the HVSR inversion of the microtremor data was then compared with the value of V_{s30} from the USGS data. The V_s value obtained is then related to depth, this is intended to determine rock sites and rock hardness levels at depth positions. The relationship between rock sites and V_s values according to the National Standardization Agency (2019) is shown in Table 2.

Table 2. Classification of rock site classes (National Standardization Agency, 2019).

Site Class	\bar{V}_{s30} (m/s)
Site Class \bar{V}_{s30} (m/s)	\bar{V}_{s30} (m/s)
SA (hard rock) > 1500	> 1500
SB (hard rock) 750 – 1500	750 – 1500
SC (hard, very dense soil, and soft rock)	350 – 750

Through the value of the amplification factor (A_0) and the dominant frequency factor (f_0) can identify the vulnerability of a soil. [24] explained that the value of through the dominant frequency factor obtained from the research area can show the nature or type of rock. The amplification factor can describe the response of a rock layer to the seismic waves that pass through it. The amplification value will increase if the wave passes through a soft material. The vulnerability value of a soil can identify the level of seismic vulnerability in a layer of soil that changes shape due to an earthquake and can trigger landslides. Based on the value of V_{s30} , the avalanche zone can be identified to a depth of 30 meters, the softer the rock, the higher the amplification value and the higher the level of vulnerability of the soil, and it is very vulnerable to causing landslides.

RESULTS AND DISCUSSION

This research was conducted in Ulu Manna District, South Bengkulu Regency, to be precise on the alternative route between Ulu Manna District and Tanjung Sakti District. This research is about "Slide Zone Based on Horizontal to Vertical Spectral Ratio Inversion in Ulu Manna District, South Bengkulu". The measurement points taken were 30 points.

The research data obtained from geophysical measurement methods, namely microtremor in the study area. The data is in the form of seismic waves, through the Geopsy software the data is processed to obtain the dominant frequency value (f_0) and amplification (A_0). Data inversion was performed using the HV-Inv software to obtain secondary or shear wave velocity values (V_s). Next, the value of V_s is calculated to obtain a value of V_{s30} . HVSR curve inversion is applied to predict the physical parameters of unknown rocks. An example of an inversion result is shown in Figure 4.

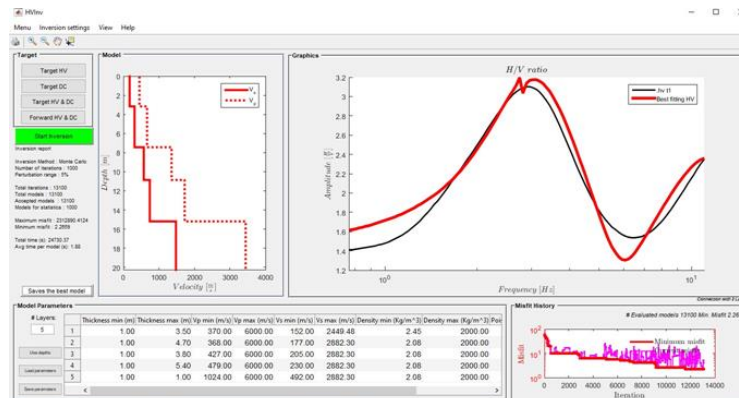


Figure 4. Graph of the results of the H/V curve inversion process at point 1.

he distribution of shear wave velocity values V_{s30} in Ulu Manna District, South Bengkulu at all microtremor measurement points, can be seen in Figure 5.

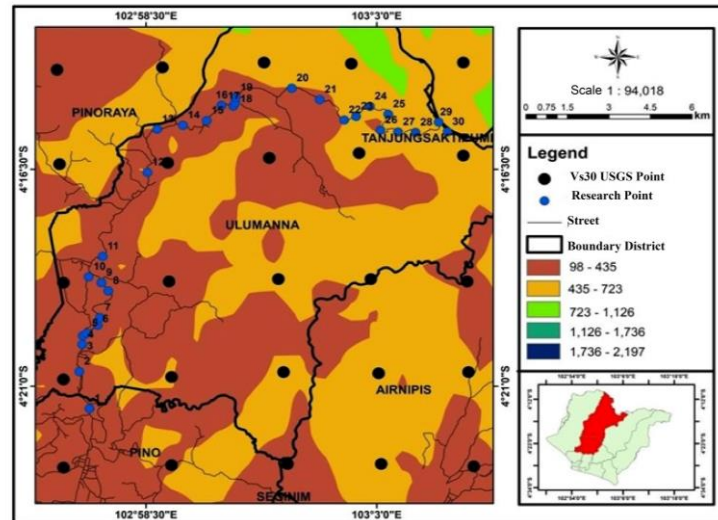


Figure 5. Map of distribution of V_{s30} values in Ulu Manna District, South Bengkulu

Shear wave velocity (V_{s30}) is a good alternative in studying and interpreting the characteristics and texture of the soil and can be used in predicting the vulnerability of an area and ensuring buildings are resistant to earthquakes [25]. The speed of shear waves decreases indicating that the soil conditions are getting softer. The speed of shear waves increases or gets bigger, describing the soil conditions in an area that are getting denser [26].

The value of the shear wave velocity (V_{s30}) obtained from measurements and calculations describes that the highest value of V_{s30} is located at point 18 at coordinates 103.0039° and -4.2518° around 766.84 m/s with site class B, namely medium rock texture. Meanwhile, the lowest shear wave velocity (V_{s30}) is at point 20 at coordinates 103.0223° and -4.2468° around 243.96 m/s with site class D, namely medium soil texture. The value of V_{s30} at each point in the research area resulting from the results of the inversion stage of the H/V curve can be seen through the distribution map of the value of V_{s30} in Figure 6.

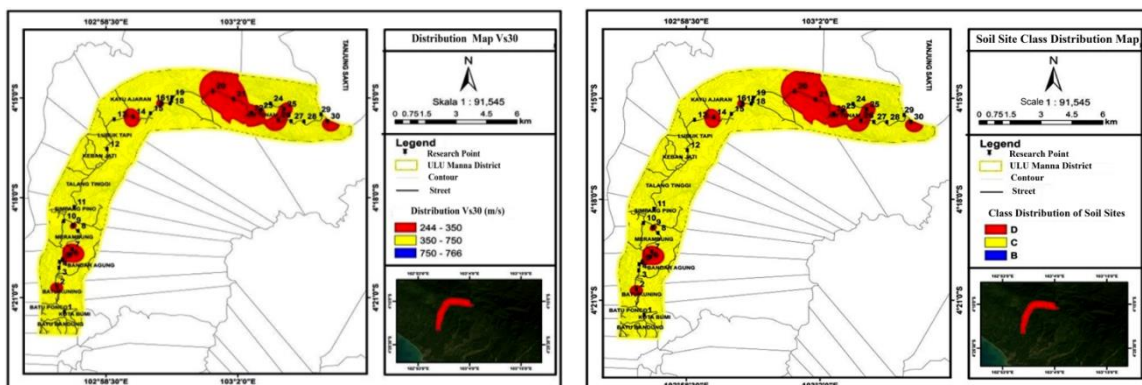


Figure 6. Map of distribution of shear wave velocity values (V_{s30}) (a) and soil site class (b).

The distribution of the V_{s30} values in Figure 6. describes that the V_{s30} value with a value range of 243.96 m/s – 347.85 m/s is symbolized in red and belongs to the site class D soil in the form of medium soil located at measurement points 2, 5, 6, 7, 9, 14, 16, 19, 20, 21, 22, 23, 25, 26, 27 and 30. The V_{s30} value ranges from 364.91 m/s – 599.62 m/s symbolized in yellow belonging to the soil class site, namely C in the form of hard soil, very dense and soft rock located at measurement points 1, 3, 4, 8, 10, 11, 12, 13, 15, 17, 24, 28 and 29. The value of V_{s30} 766.84 m/s is symbolized in dark blue belonging to site class, namely B in the form of rock being at measurement point 18. A low value of V_{s30} belongs to zone of high seismic susceptibility with a thick layer of sediment is located in class D. Meanwhile, classes B and C are classified as zones of low seismic susceptibility to the danger of earthquake natural disasters which can cause landslides.

Based on the geological map of South Bengkulu Regency, the geological formations in Ulu Manna District are dominated by the Lemau Formation, Seblat Formation and Simpang Aur Formation (BPBD South Bengkulu, 2021). The Lemau, Seblat and Simpang Aur Formations are composed of claystone, sandstone, silt, conglomerate and mudstone consisting of tufan sandstone and mollusk shells [27]. Claystone, silt and rocks or soil containing sand have physical properties that are very soft in consistency, very high in compressibility, namely 0.38 and low carrying capacity between soils which causes the potential for settlement and landslides [28]. There is a difference in the value of V_s30 from the results of the microtremor data field and the USGS data, but the value is not much different. This is because the data V_s30 from USGS only interprets through the slope parameter obtained from satellite imagery [29]. So, it can be concluded that the difference in the value of V_s30 occurs due to changes in the slope caused by geological conditions.

Based on BMKG data from Bengkulu Province for the last 10 years (2013-2022) it was found that South Bengkulu, to be precise, in Ulu Manna District has very high rainfall with a total rainfall reaching 3854 – 5513 mm per year (BMKG Bengkulu Province, 2023), which can be seen on . Very high rainfall can cause the potential for landslides to occur. The Ulu Manna sub-district also has a shear strain on the soil (GSS) value of the order 10^{-4} to 10^{-3} , in which the soil properties are elastic-plastic dynamics. The elastic-plastic condition is classified as one of the characteristics of an area having a high potential for landslides [30].

The results of the research on microtremor data, especially from the V_s30 value and rainfall data, can be concluded that Ulu Manna District is classified as a zone of high seismic vulnerability found at points included in site class D and low seismic vulnerability zones located at points belonging to site class B and C. Zones of low seismic vulnerability are safer than zones of higher seismic vulnerability which are more prone to seismic disturbances or earthquake disasters and can cause landslides.

CONCLUSION

Based on the results of research that has been conducted regarding "Slide Zone Based on Horizontal to Vertical Spectral Ratio Inversion in Ulu Manna District, South Bengkulu", it can be concluded that:

The shear wave velocity (V_s30) in Ulu Manna District, South Bengkulu Regency ranges from 243.96 m/s – 766.84 m/s. The largest value of V_s30 is at coordinates 103.0039° and -4.2518° around 1308.69 m/s and the lowest value of V_s30 is at coordinates 103.0223° and -4.2468° around 136.76 m/s. The results of the inversion of the HVSr curve are dominated by site class D soils with moderate soil texture. Site class classification based on the value of V_s30 in the research area belongs to class B, C and D.

Classification of low seismic vulnerability zones in Ulu Manna District is at points belonging to site class B and C soils. Zones of high seismic vulnerability are at points belonging to class D. The V_s30 value from the USGS has soil classifications C and D. From the results of inversion of microtremor data with data from USGS, there is a difference in the value of V_s30 . The ratio of V_s30 between microtremor data and USGS ranges from 0.56 – 1.76.

The results of rainfall data (2013 – 2022) in Ulu Manna District have very high rainfall with a total of 3854 – 5513 mm per year. Based on the results of the study coupled with rainfall data, there are areas that are prone to potential landslides in Ulu Manna District, namely points belonging to class D and E.

ACKNOWLEDGEMENT

The author would like to thank the Geophysics Laboratory, Department of Physics, Faculty of Mathematics and Natural Sciences, University of Bengkulu, which has assisted in field measurements and laboratory tests. This study is also supported by students, especially in data acquisition. The researcher also thanks the related units, such as the South Bengkulu Regional Government, which gave permission for this research.

REFERENCES

- [1] F. Humairah, Zona Longsoran Berdasarkan Inversi Horizontal to Vertical Spectral Ratiodi Kecamatan Ulu Manna, Bengkulu Selatan, Skripsi, Prodi Geofisika, 2023.
- [2] A. Ambikapathy, J. K Catherine, V. K Gahalaut*, M Narsaiah, A. Bansal and P. Mahesh, Bengkulu earthquake, its rupture model and implications for seismic hazard. Earth Syst. Sci. 119, No. 4, August 2010, pp. 553–560.

- [3] A. I. Hadi, Refrizon, M. Farid, B. Harlianto, and J. I. Sari, Landslide Potential Investigation for Disaster Risk Reduction in Central Bengkulu Regency, Bengkulu Province, Indonesia, *Indonesian Journal on Geoscience*, 8 (3), p.313-328. 2021.
- [4] K. Sassa., H. Fukuoka., G. Scarascia-Mugnozza, G., S. Evans, Earthquake-induced landslides: Distribution, motion, and mechanisms. Special Issue for the Great Hanshin earthquake disaster, *Soils and Foundations*, (1996). 53–64.
- [5] J. Delgado., C. L. Casado., A. Esteves., J. Giner., A. Cuenca., S. Molina.,.. Mapping soft soils in the Segura river valley (SE Spain): a case study of microtremors. *Journal Applied Geophysics*, (2000) 45: 19-32.
- [6] K. Chousianitis V. D. Gaudio, I. Kalogeras, A. Ganas, Predictive model of Arias intensity and Newmark displacement for regional scale evaluation of earthquake-induced landslide hazard in Greece, 2014.
- [7] M. Farid, A. I. Hadi, Measurement of Shear Strain in Map Liquefaction Area for Earthquake Mitigation in Bengkulu City. *TELKOMNIKA*, Vol.16, No.4, August 2018, pp. 1597~1606.
- [8] R. K. Nepop, A. R. Agatova., Estimating Magnitudes of Prehistoric Earthquakes from Landslide Data: First Experience in Southern Altai. *Russian Geology and Geophysics*, (2008) 49, p.144-151, Elsevier. Amsterdam, The Netherlands DOI: 10.1016/j.ragg.2007.06.013.
- [9] M. Noorwantoro., R. Asmaranto., D. Harisuseno., Analisa Kawasan Rawan Bencana Tanah Longsor di DAS Upper Brantas Menggunakan Sistem Informasi Geografi", Universitas Brawijaya, Malang. (2013).
- [10] B. Yulihanto, B., Situmorang, A. Nurdjajadi., B. Sain, tructural Analysis of The Onshore Bengkulu Forearc Basin and Its Implication for Future Hydrocarbon Exploration Activity", *Proceedings 24th Annual Convention: IPA*, pp. 85-86 (1995).
- [11] D. H. Santoso, Suharwanto, T. Prasetyo, Analisis Kestabilan Lereng dan Pengelolaan Lereng Akibat Penambangan Andesit di Sebagian Kecamatan Bagelan Purworejo , *Jurnal Geografi*, 18(1) (2021) 46-51.
- [12] V. Pazzi, S. Morelli, R. Fanti, Review Article A Review of the Advantages and Limitations of Geophysical Investigations in Landslide Studies. *Hindawi International Journal of Geophysics Volume 2019*, Article ID 2983087, 27 pages <https://doi.org/10.1155/2019/2983087>.
- [13] R. Nura, E. Hartantyo. Penentuan Area Potensi Tanah Longsor Berdasarkan Analisis Mikrotremor di Dusun Jeruk dan Sekitarnya Kecamatan Samigaluh Kabupaten Kulon Progo. *POSITRON Vol. 11, No. 2* (2021), Hal. 77 – 85.
- [14] Y. Nakamura, „Clear Identification of Fundamental Idea of Nakamura’s Technique and Its Application”, *Proceeding XII World Conference of Earthquake Engineering*, Tokyo University, Japan (2000).
- [15] E. P. Utomo, Longsoran yang Dipicu Gempa Bumi (Studi Kasus: Longsoran Warungkiara, Sukabumi, Jawa Barat dan Longsoran Raksasa Gowa, Sulawesi Selatan), *Prosiding Pemaparan Hasil Penelitian Puslit Geoteknologi 2008*, 10 Desember 2008. ISBN: 978-979-8636-15-8.
- [16] D. I. Sutiyono., S. Balamba., A. N. Sarajar, Analisis Stabilitas Lereng Akibat Gempa di Ruas Jalan Noongan – Pangu", *Tekno*, 15(67). IS(2017SN: 0215- 9617.
- [17] W. A. Prakoso, I. N. Sukanta, Evaluation of V_s30 Estimating Models for Indonesia", *Jurnal Teknologi (Sciences & Engineering)*, 77 (11), (2015) pp. 83-90.
- [18] R. Anbalagan, (1992) „Landslide Hazard Evaluation and Zonation, Mapping in Mountain Terrain", *Engineering Geology*, 32, pp. 269-277. (2007).
- [19] L. F. Aristizábal, ., J. Mauricio., E. B. Alex ., I. T. Héctor ., & P. A. Steven . 2015. Monitoring Coffee berry borer, *Hypothenemus hampei* (Coleoptera: Curculionidae) populations with alcohol-baited funnel traps in coffee farms in Colombia. *Journal of Florida Entomologist*, 98(1): 380-383.
- [20] B. G. M. Saka, R. O. Tarru, E. Lolang, A. Pakiding., Identification of Slip Area in Makale Selatan District Using the Geoelectric Method , *Jurnal Geoelebes Vol. 7No. 1*, April 2023, 77–88.
- [21] X. Wang, L. D. Zilio, J. K. Morgan, D. S. Kammer, Non-Precursory Accelerating Aseismic Slip During Rupture Nucleation. *Journal of Geophysical Research: Solid Earth*, *Journal of Geophysical Research: Solid Earth*, 2023.
- [22] S. K. Singh., M. Ordaz., J. F. Pacheco, Advanced in Seismology with Impact on Earthquake. *International Handbook of Earthquake and Engineering Seismology*, Volume 81 (2003).
- [23] S. H. Soenarmo, I. A. Sadisun, E. Saptohartono., Kajian Awal Pengaruh Intensitas Curah Hujan Terhadap Pendugaan Potensi Tanah Longsor Berbasis Spasial di Kabupaten Bandung, Jawa Bara, *Jurnal Geoaplika* (2008) Volume 3, Nomor 3, hal. 133 – 141 133.
- [24] Y. Nakamura. A method for dynamic characteristic estimation of subsurface using microtremor on the ground surface. *Q.R. of R.T.I.* (1989) 30-1, p. 25-33.

-
-
- [25] C. B. Park., R. D. Miller., J. Xia. Multichannel Analysis of Surface Waves (MASW): A Summary Report of Technical Aspects, Experimental Result, and Perspective, Kansas Geological Survey (1997).
- [26] J. Milsom, A. Eriksen, Field Geophysics. Four Edition. New Delhi: John Wiley dan Sons (2011).
- [27] B. Yulihanto, B., Situmorang, A. Nurdjajadi., B. Sain, Structural Analysis of The Onshore Bengkulu Forearc Basin and Its Implication for Future Hydrocarbon Exploration Activity“, Proceedings 24th Annual Convention: IPA, pp. 85-86 (1995).
- [28] Wardoyo, Sarwondo, D. Destiasari, Wahyudin, Wiyono, G. Hasibuan, W. P. Sollar, „Atlas Sebaran Tanah Lunak Indonesia Edisi 2019“, Badan Geologi Kementerian Energi dan Sumber Daya Mineral. ISBN: 978-602-9105-80-3. M. P. Brown and K. Austin, The New Physique (Publisher Name, Publisher City, 2005), pp. 25–30.
- [29] C. Athanasius., dan A. Solikhin, Pendugaan Kecepatan Gelombang Permukaan (VS30) di Pulau Sulawesi Berdasarkan Klasifikasi Geomorfologi dan Aplikasinya“, Badan geologi Kementerian Energi dan Sumber Daya Mineral, pp. 1-8, Kemenerian ESDM, Jakarta (2019).
- [30] I. Awaliyah, A. I. Hadi, M. Farid, D. I. Fadli., A.J. Akbar, (2023) „Microzonation Site Effects and Shear Strain during Earthquake Induced Landslide Using HVSR Measurement in Ulu Mana Sub-District, South Bengkulu Regency Indonesia“, Jurnal Penelitian Pendidikan IPA, 9(2), pp. 592-599, Universitas Mataram, NTB.

Application of Quadratic Programming Using Wolfe Method on Portfolio Optimization

Syaripuddin^{*(1)}, Fidia Deny Tisna Amijaya⁽¹⁾, Wasono⁽¹⁾, Shanaz Tulzahrah⁽¹⁾,
Rara Suciati⁽¹⁾

⁽¹⁾Department of Mathematics, Faculty of Mathematics and Natural Sciences,
Mulawarman University, Indonesia,

*Corresponding author: syarifrahman2014@gmail.com

Abstract. Quadratic programming (QP) can be applied to optimization problems, one of which is the stock portfolio optimization problem. The stock portfolio can be modeled into a QP-model using Markowitz mean-variance method developed by Harry Markowitz (1956). This study aims to determine the optimal results on portfolio problems using the Wolfe method. The data is collected from 10 stock companies that paid the highest dividends in the IDX High Dividend 20 category for the 2022 period. The results of this study focus on discussing two portfolio models. First, The portfolio that has negative covariance with a significant difference in expected return rates. Based on Markowitz mean-variance, the expected profit rate is 3.432% and the risk is 0.505%. The optimal proportion of shares invested for PT. Astra International Tbk is 76.401% and for PT. Indo Tambangraya Megah Tbk is 23.599%. Second, The portfolio that has a negative covariance with a slight difference in the expected return rates. Based on the mean variance of Markowitz, the expected profit rate is 2.471% and the risk is 0.141%. The optimal proportion of shares invested for PT. Bank Rakyat Indonesia (Persero) Tbk is 56.089% and for PT. Mitra Pinasthika Mustika Tbk is 43.911%.

Keywords: Markowitz Mean-Variance, Optimal Portfolio, Quadratic Programming, Wolfe Method

INTRODUCTION

Quadratic programming is a special form of nonlinear programming which has special characteristics, namely having quadratic objective function and linear constraint function(s) [1]. Even though quadratic programming is part of nonlinear programming, its solution still adopts several solving methods used in linear programming problems. One interesting method to discuss is the Wolfe method, this method transforms a quadratic programming problem into a linear programming problem. Wolfe (1959) modified the simplex method to solve quadratic programming problems by adding the Karush-Kuhn-Tucker (KKT) condition and changing the shape of the quadratic objective function to a linear function [2].

The problem of portfolio optimization is motivated by the desire of investors to invest in risky assets (stocks), with the hope of still obtaining optimal returns in accordance with the risks they are willing to bear. Therefore, investors can diversify by trying to combine several securities to obtain returns and minimize risk, in other words forming a portfolio. A stock portfolio is a combination of several shares owned by an investor or a particular entity. A stock portfolio can be an alternative for making investment decisions for potential investors [3]. The stock portfolio can be modeled into a quadratic programming model using the mean variance Markowitz method [4].

Several studies that have discussed quadratic programming are [5] which discusses methods of solving quadratic programming problems, [6] discusses solving stock portfolio optimization problems using the application of the Kuhn-Tucker method, [7] discusses the Frank-Wolfe type algorithm to solve Mean-Risk Optimization, [8] discusses the application of the Wolfe method to a case study of stock portfolio optimization. Based on this, the present research discusses in more detail by providing examples and case studies regarding the application of the Wolfe method to quadratic programming models to be applied as an optimal portfolio selection model.

METHODS

Quadratic Programming

The general form of a quadratic programming problem is defined as follows [9]:

$$\text{Maximize } Z = f(x) = \sum_{j=1}^n c_j x_j + \frac{1}{2} \sum_{j=1}^n \sum_{k=1}^n q_{jk} x_j x_k \quad (1)$$

with constraints

$$\sum_{j=1}^n c_j x_j \leq b_i, \quad i = 1, 2, \dots, m$$

$$x_i \geq 0, \quad j = 1, 2, \dots, n$$

where $q_{jk} = q_{kj}$ for each $j = 1, 2, \dots, n$, $k = 1, 2, \dots, n$ and $b_i \geq 0$ for each $i = 1, 2, \dots, m$, and the quadratic form $\sum_{j=1}^n \sum_{k=1}^n q_{jk} x_j x_k$ is negative semi definite.

Wolfe's Method

The Wolfe method was developed by Wolfe in 1959, this method is used to solve quadratic programming problems. The stages of the Wolfe method for quadratic programming problems involve the Lagrange method and the Karush-Kuhn-Tucker (KKT) condition. The steps in the Wolfe method are presented as follows:

1. Determine the objective function in the form of non-linear equations and constraints in the form of linear inequalities.
2. Form the Karush Kuhn Tucker conditions according to the objective function and constraints obtained.
3. Determine the condition of complementary slackness.
4. Add artificial variables for Karush Kuhn Tucker conditions that do not yet have a base variable
5. Define new objective function.
6. Calculate a two-stage simplex iteration with the Wolfe method.
7. Substitute the results obtained from the simplex iteration process into the initial objective function, so that optimal results are obtained.

Based on the Wolfe method algorithm, Equation (1) can be modeled into a linear programming model as follows:

$$\text{Minimize } Z_v = v_1 + v_2 + \dots + v_n \quad (2)$$

with constraints

$$-\sum_{k=1}^n q_{jk} x_k + \sum_{i=1}^m a_{ij} \lambda_i - \mu_j + v_j \leq c_i, \quad j = 1, 2, \dots, n$$

$$\sum_{k=1}^n a_{ij} x_j + s_i = b_i, \quad i = 1, 2, \dots, m$$

$$v_i, \lambda_i, \mu_i, x_i \geq 0, \quad j = 1, 2, \dots, n, i = 1, 2, \dots, n$$

and satisfy the complementary requirements

$$\sum_{j=1}^n \mu_j x_j + \sum_{i=1}^m \lambda_i s_i = 0$$

where $\lambda_i s_i = 0$ and $\mu_j x_j = 0$, ($i = 1, 2, \dots, m$ and $j = 1, 2, \dots, n$).

Additional rules for Wolfe's method when selecting the basis variable are as follows:

1. If x_i has been selected as the basis variable, then its complement, namely μ_i , cannot be selected as the basis variable
2. If s_i has been selected as the basis variable, then its complement, namely λ_i , cannot be selected as the basis variable.

Portfolio Theory

Portfolio is an investment that consists of various stocks so that an efficient combination can be made so that investors can get the desired rate of return with minimal risk. The steps in determining the optimal portfolio using the Markowitz model are as follows [10]:

1. Calculating Individual Returns

$$R_{it} = \frac{P_{it} - P_{i(t-1)} + \left(\frac{D}{12}\right)}{P_{i(t-1)}} \quad (3)$$

where

- R_{it} : individual realization returns
 P_{it} : closing price of the i -th stock in the t -th period
 $P_{i(t-1)}$: closing price of the i -th stock in the $(t-1)$ -th period
 D : value of the i -th stock dividend

2. Calculation the Individual Return Expectations

$$E(R_i) = \frac{\sum_{i=1}^N R_{it}}{N} \quad (4)$$

where

- $E(R_i)$: expected return on i -th asset
 N : number of returns that occur in the observation period
 R_p : portfolio realization return

3. Calculating Individual Risk

$$\sigma_i^2 = \frac{\sum_{j=1}^N (R_{it} - (E(R_i)))^2}{N - 1} \quad (5)$$

- σ_i^2 : variance of investment in stock i
 $E(R_i)$: expected return on assets i
 R_{it} : return of the i -th asset in the t -th period
 N : number of returns that occur in the observation period

4. Calculating the Covariance

$$cov R_i R_j = \frac{\sum_{i,j=1}^N [R_{it} - \sum_{i=1}^N E(R_i)][R_{jt} - \sum_{j=1}^N E(R_j)]}{N - 1} \quad (6)$$

where

- σ_i^2 : variance of investment in stock i
 σ_p^2 : portfolio risk
 $E(R_i)$: expected return on i -th asset
 R_{it} : return of the i -th asset in the t -th period
 N : number of returns that occur in the observation period
 $cov R_i R_j$: covariance of return on the i -th and j -th assets

5. Determine the proportion of funds from each stock portfolio using a quadratic programming model based on the mean variance Markowitz model. The chosen model is to minimize risk with a certain rate of return given as follows:

Minimize

$$Z = \sum_{i=1}^N \sum_{j=1}^N x_i x_j \sigma_{ij} \quad (7)$$

with constraints:

$$\sum_{i=1}^N x_i = 1$$

$$\sum_{i=1}^N E(R_i) x_i \geq E(\hat{R}_i)$$

$$x_i \geq 0, i = 1, \dots, N$$

where

- Z : objective function
 $E(R_i)$: individual expected return
 $E(\hat{R}_i)$: individual expected return minimum
 σ_i^2 : individual stock risk
 σ_{ij} : portfolio stock risk
 N : the number of returns that occur in the observation period

Data

This study uses secondary data in the form of time series data, namely company stock data obtained from the website of PT. Indonesia Stock Exchange, namely on the page <https://www.idx.co.id/id>. The variables used are the adjusted closing value and the dividend value of each share per time period, namely January 2022 - December 2022. The adjusted closing value is the closing value of the shares in a certain period. Dividends are cash payments made by the company to shareholders.

RESULTS AND DISCUSSION

Description of Research Data

The research data used is stock data of 10 companies in Indonesia that pay out the largest dividends in the IDX High Dividend 20 category in January 2022 – December 2022. The IDX High Dividend 20 is an index that measures the price performance of 20 stocks that have distributed cash dividends for the last 3 years and have dividends tall one. The following is a list of shares of 10 companies in Indonesia that distribute the largest dividends in the IDX High Dividend 20 throughout 2022:

1. ADMF: Adira Dinamika Multi Finance Tbk.
2. ASII: Astra International Tbk.
3. BBRI: Bank Rakyat Indonesia (Persero) Tbk.
4. BMRI: Bank Mandiri (Persero) Tbk.
5. HEXA: Hexindo Adiperkasa Tbk.
6. INDF: Indofood Sukses Makmur Tbk.
7. ITMG: Indo Tambangraya Megah Tbk.
8. MPMX: Mitra Pinasthika Mustika Tbk.
9. PTBA: Bukit Asam Tbk.
10. UNTR: United Tractors Tbk.

The return value of 10 companies obtained using Equations (3) and (4) can be seen in Table 1.

Table 1. Individual Realization Return Value

<i>Periode</i>	ADMF	ASII	BBRI	BMRI	HEXA	INDF	ITMG	MPMX	PTBA	UNTR
January	0	0	0	0	0	0	0	0	0	0
February	0.04333	0.06383	0.12169	0.03431	0.06189	-0.01595	0.26646	-0.01005	0.12541	0.08332
March	0.09579	0.13784	0.02754	0.03006	0.24753	-0.03643	0.09817	0.17526	0.06923	0.03220
April	-0.02282	0.15581	0.08977	0.19091	0.00003	0.06288	0.02080	-0.02222	0.18158	0.19087
May	0.07049	-0.02648	-0.04626	-0.04692	0.05765	0.05145	0.40388	0.01382	0.20351	0.07148
June	-0.02114	-0.07162	-0.10050	-0.06411	-0.05509	0.07184	-0.10456	0.01382	-0.14185	-0.08795
July	0.02499	-0.04169	0.05415	0.04796	0.05067	-0.03204	0.31264	-0.07834	0.33993	0.14250
August	0.00313	0.10653	-0.00122	0.07312	0.04069	-0.08101	0.00923	0.05076	0.00172	0.05254
September	0.02461	-0.04676	0.03795	0.06837	0.03956	0.01238	0.07100	-0.02451	-0.00532	-0.02593
October	0.04224	0.00737	0.03891	0.12255	0.02802	0.07438	0.10357	0.06122	-0.04859	-0.01151
November	0.05821	-0.07421	0.07413	0.00048	-0.04845	0.00359	-0.05848	0.13659	-0.01346	-0.01823
December	0.00006	-0.05397	-0.00508	-0.05415	-0.02880	0.04623	0.05712	-0.01304	-0.01385	-0.14875
Expected Return	0.02657	0.01305	0.02426	0.03355	0.03281	0.01311	0.09832	0.02528	0.05819	0.02338
Individual Risk	0.00127	0.00656	0.00274	0.00550	0.00603	0.00234	0.02016	0.00491	0.01680	0.00844

The covariance value between shares of 10 companies obtained using Equations (5) and (6) can be seen in Table 2.

Table 2 Share Portfolio Covariance Value

Company	ADMF	ASII	BBRI	BMRI	HEXA	INDF	ITMG	MPMX	PTBA
ASII	0.00041								
BBRI	0.00452	0.01914							
BMRI	-0.00644	0.03579	0.02902						
HEXA	0.02040	0.03664	0.00687	0.00586					
INDF	-0.00429	-0.01071	-0.00690	-0.00050	-0.01891				
ITMG	0.02937	-0.0087	0.02932	0.01379	0.07236	-0.00939			
MPMX	0.01669	0.01927	-0.00301	-0.00896	0.02189	-0.00791	-0.04126		
PTBA	0.01051	0.02094	0.03525	0.02824	0.03918	-0.02122	0.20212	-0.04241	
UNTR	-0.00084	0.04090	0.02373	0.03172	0.01318	-0.01394	0.02189	-0.02596	0.09809

Calculating Return and Risk Expected Portfolio

Expected return is the expected return of the investment made. The portfolio is composed of two companies. If it is assumed that each company obtains the same proportion of 50%, then the expected return value of each stock portfolio can be determined.

Expected return of PT. Astra International Tbk and PT. Indo Tambangraya Megah Tbk portfolio:

$$\begin{aligned}
 E(R_{ASII\&ITMG}) &= \sum_{i=1}^N (k_i E(R_i)) \\
 &= (50\% \cdot 1,305\%) + (50\% \cdot 9,832\%) \\
 &= (0,5 \cdot 0,01305) + (0,5 \cdot 0,09832) \\
 &= 0,05569 = 5,569\%
 \end{aligned} \tag{8}$$

Expected return of PT. Bank Rakyat Indonesia (Persero) Tbk and PT. Mitra Pinasthika Mustika Tbk portfolio:

$$\begin{aligned}
 E(R_{BBRI\&MPMX}) &= \sum_{i=1}^N (k_i E(R_i)) \\
 &= (50\% \cdot 2,426\%) + (50\% \cdot 2,528\%) \\
 &= (0,5 \cdot 0,02426) + (0,5 \cdot 0,02528) \\
 &= 0,02477 = 2,477\%
 \end{aligned} \tag{9}$$

Calculating Portfolio Risk

Portfolio risk is the variance of returns from the securities that make up the portfolio so that the portfolio risk may be less than the weighted average risk of each single security.

Portfolio risk of investment on PT. Astra International Tbk and PT. Indo Tambangraya Megah Tbk:

$$\begin{aligned}
 \sigma_{ASII\&ITMG}^2 &= \sum_{i=1}^N k_i^2 \sigma_i^2 + \sum_{j=1}^N k_j^2 \sigma_j^2 + 2 \sum_{i=1}^N \sum_{j=1}^N k_i k_j \sigma_{ij} \\
 &= ((50\%)^2 \cdot 0,681\%) + ((50\%)^2 \cdot 2,358\%) \\
 &\quad + (50\% \cdot 50\% \cdot (-0,01374\%)) \\
 &= (0,5^2 \cdot 0,00681) + (0,5^2 \cdot 0,02358) + (0,5 \cdot 0,5 \cdot (-0,0001374)) \\
 &= 0,00725 = 0,725\%
 \end{aligned} \tag{10}$$

Portfolio risk of investment on PT. Bank Rakyat Indonesia (Persero) Tbk dan PT. Mitra Pinasthika Mustika Tbk:

$$\begin{aligned}
 \sigma_{BBRI\&MPMX}^2 &= \sum_{i=1}^N k_i^2 \sigma_i^2 + \sum_{j=1}^N k_j^2 \sigma_j^2 + 2 \sum_{i=1}^N \sum_{j=1}^N k_i k_j \sigma_{ij} \\
 &= ((50\%)^2 \cdot 0,366\%) + ((50\%)^2 \cdot 0,508\%) + (50\% \cdot 50\% \\
 &\quad \cdot (-0,292\%)) \\
 &= (0,5^2 \cdot 0,00366) + (0,5^2 \cdot 0,00508) + (0,5 \cdot 0,5 \cdot (-0,0029)) \\
 &= 0,00141 = 0,141\%
 \end{aligned} \tag{11}$$

Quadratic Programming Model Based on the Mean Variance Markowitz Model

This study focuses on discussing two stock portfolio models, namely (1) Model A with the characteristics of a negative covariance stock portfolio with a much different expected rate of return, namely the stock portfolio of PT. Astra International Tbk and PT. Indo Tambangraya Megah Tbk. (2) Model B with the characteristics of a negative covariance stock portfolio with a difference in the expected rate of return that is not much different, namely PT. Bank Rakyat Indonesia (Persero) Tbk and PT. Mitra Pinasthika Mustika Tbk. In the following, a quadratic programming model is presented for both stock portfolios:

1. Model A

The model A stock portfolio is a portfolio with a negative covariance with significantly different expected return rates, namely ASII and ITMG, where the covariance value of the two stocks is -0.0087 with the expected return from ASII of 0.01305 and the expected return of ITMG of 0.09832. In addition, there is a risk value from ASII of 0.00681 and a risk value of ITMG of 0.02358. Therefore, the quadratic programming model is obtained as follows:

Minimize

$$Z = 0.00681x_1^2 + 0.02358x_2^2 - 0.00137x_1x_2 \quad (12)$$

with constraints:

$$\begin{aligned} x_1 + x_2 &= 1 \\ 0.01305x_1 + 0.09832x_2 &\leq 0.01305 \\ x_1, x_2 &\geq 0 \end{aligned}$$

The quadratic programming model in Equation (12) is transformed into a linear programming model using the Wolfe method as follows:

Minimize

$$Z = z_1 + z_2 + z_3 + z_4 \quad (13)$$

with constraints:

$$\begin{aligned} 0.01362x_1 - 0.00137x_2 + \lambda_1 - \lambda_2 - 0.01305\lambda_3 - \mu_1 + z_1 &= 0 \\ -0.00137x_1 + 0.04716x_2 + \lambda_1 - \lambda_2 - 0.01305\lambda_3 - \mu_1 + z_1 &= 0 \\ x_1 + x_2 + y_1 &= 1 \\ x_1 + x_2 - y_1 + z_3 &= 1 \\ 0.01305x_1 + 0.09832x_2 - y_3 + z_4 &= 0.01305 \end{aligned}$$

and satisfy the complementary requirements

$$\lambda_1 y_1 = 0, \lambda_2 y_2 = 0, \lambda_3 y_3 = 0, \mu_1 x_1 = 0, \mu_2 x_2 = 0$$

where

$$x_1, x_2, y_1, y_2, \lambda_1, \lambda_2, \lambda_3, \mu_1, \mu_2 \geq 0$$

The linear programming model in Equation (8) is solved using the Wolfe method. The Wolfe method iteration process is stated in the following table:

Table 3. Iteration of the Wolfe method on the linear programming model

Iteration	Basis Variabel	Right side Value	Entering Variable	Leaving Variable
1	$\begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \\ y_1 \end{bmatrix}$	1.01305	Select x_2 as the entering variable	Select z_2 as the leaving variable
2	$\begin{bmatrix} z_1 \\ x_2 \\ z_3 \\ z_4 \\ y_1 \end{bmatrix}$	1.01305	Select μ_2 as the entering variable	Select z_1 as the leaving variable
3	$\begin{bmatrix} x_1 \\ x_2 \\ z_3 \\ z_4 \\ y_1 \end{bmatrix}$	1.01305	Select x_1 as the entering variable	Select μ_2 as the leaving variable
4	$\begin{bmatrix} x_1 \\ x_2 \\ z_3 \\ z_4 \\ y_1 \end{bmatrix}$	1.01305	Select λ_2 as the entering variable	Select z_4 as the leaving variable

Table 4. Iteration of the Wolfe method on the linear programming model (cont.)

Iteration	Basis Variabel	Right side Value	Entering Variable	Leaving Variable
5	$\begin{bmatrix} x_1 \\ x_2 \\ z_3 \\ \lambda_2 \\ y_1 \end{bmatrix}$	0.60660	Select y_3 as the entering variable	Select z_3 as the leaving variable
6	$\begin{bmatrix} x_1 \\ x_2 \\ y_3 \\ \lambda_2 \\ y_1 \end{bmatrix}$	0	Optimal	Optimal

Then, the solution to the linear programming problem is obtained as follows:

$$\begin{bmatrix} x_1 \\ x_2 \\ y_3 \\ \lambda_2 \\ y_1 \end{bmatrix} = \begin{bmatrix} 0.76401 \\ 0.23598 \\ 0.02012 \\ 0.01008 \\ 0 \end{bmatrix}$$

The optimal proportion of ASII is 0.76401 or 76.401% and the optimal proportion of ITMG is 0.23598 or 23.598%.

2. Model B

The model B stock portfolio is a portfolio with a negative covariance with a slight difference in expected return rates, namely BBRI and MPMX, where the covariance value of the two stocks is -0.00301 with the expected return value from BBRI companies of 0.02426 which will be the coefficient of the variable x_1 and the value the expected return from the MPMX company of 0.02528 will be the coefficient of the variable x_2 . In addition, there is a risk value for BBRI companies of 0.00366 and a risk value for MPMX companies of 0.00508. Therefore, the form of the quadratic programming model is obtained as follows:

Minimize

$$Z = 0.00366x_1^2 + 0.00508x_2^2 - 0.00292x_1x_2 \quad (14)$$

with constraints:

$$\begin{aligned} x_1 + x_2 &= 1 \\ 0.02426x_1 + 0.02528x_2 &\leq 0.02426 \\ x_1, x_2 &\geq 0 \end{aligned}$$

The quadratic programming model in Equation (14) is transformed into a linear programming model using the Wolfe method as follows:

Minimize

$$Z = z_1 + z_2 + z_3 + z_4 \quad (15)$$

with constraints:

$$\begin{aligned} 0.00732x_1 - 0.00292x_2 + \lambda_1 - \lambda_2 - 0.02426\lambda_3 - \mu_1 + z_1 &= 0 \\ -0.00292x_1 + 0.001016x_2 + \lambda_1 - \lambda_2 - 0.02528\lambda_3 - \mu_1 + z_1 &= 0 \\ x_1 + x_2 + y_1 &= 1 \\ x_1 + x_2 - y_1 + z_3 &= 1 \\ 0.02426x_1 + 0.0258x_2 - y_3 + z_4 &= 0.02426 \end{aligned}$$

and satisfy the complementary requirements

$$\lambda_1 y_1 = 0, \lambda_2 y_2 = 0, \lambda_3 y_3 = 0, \mu_1 x_1 = 0, \mu_2 x_2 = 0$$

where

$$x_1, x_2, y_1, y_2, \lambda_1, \lambda_2, \lambda_3, \mu_1, \mu_2 \geq 0$$

The linear programming model in Equation (15) is solved using the Wolfe method. Thus, the solution to the linear programming problem is obtained as follows:

$$\begin{bmatrix} x_1 \\ x_2 \\ y_3 \\ \lambda_2 \\ y_1 \end{bmatrix} = \begin{bmatrix} 0.56089 \\ 0.43908 \\ 0.00044 \\ 0.00028 \\ -0.027118 \end{bmatrix}$$

The optimal proportion of BBRI is 0.56089 or 56.089% and the optimal proportion of MPMX is 0.43908 or 43.908%.

CONCLUSION

The result of stock portfolio optimization based on quadratic programming using the Wolfe method portfolio shows that investing optimal proportion in PT Astra International Tbk by 76.401% and PT. Indo Tambangraya Megah Tbk by 23.599% has the expected profit rate of 3.432% and the risk of 0.505%. Meanwhile investing optimal proportion in PT. Bank Rakyat Indonesia (Persero) Tbk by 56.089% and PT. Mitra Pinasthika Mustika Tbk by 43.911% has the expected profit rate of 2.471% and the risk of 0.141%

ACKNOWLEDGEMENT

The authors gratefully acknowledge the funding of Faculty of Mathematics and Natural Sciences, Mulawarman University Samarinda Indonesia in 2023. [1410/UN17.7/PP/2023]

REFERENCES

- [1] F.S. Hiller, J. Gerald, & Lieberman, Introduction to Operations Research, (The McGraw-Hill Companies, New York, 2021)
- [2] Syaripuddin, H. Suprajitno, & Fatmawati, Extension of Wolfe Method for Solving Quadratic Programming with Interval Coefficients, (Hindawi, London, 2017), pp. 1-6
- [3] L. Cheng, M. Shadabfar, & A.S. Khojine, A State-of-the-Art Review of Probabilistic Portfolio: Management for Future Stock Markets, (mathematics, Basel, 2023), pp. 1 – 34
- [4] N.A. Halim & A. Yuliati, Markowitz Model Investment Portfolio Optimization: a Review Theory, (International Journal of Research in Community Service, Indonesia, 2020), pp. 14 – 18
- [5] W. Liu, L. Yang, & B. Yu, A Lifting-Penalty Method for Quadratic Programming with a Quadratic Matrix Inequality Constraint, (mathematics, Basel, 2023), pp. 1 – 11
- [6] Z. Dai, A Closer Look at the Minimum-Variance Portfolio Optimization Model, (Hindawi, London, 2019), pp. 1 – 8
- [7] C. Duchheim, M. De Santis, F. Rinaldi, & L. Trieu, A Frank-Wolfe Based Branch-and-Bound Algorithm for Mean-Risk Optimization, (arXiv:1507.05914v4, 2017), pp. 1 – 22
- [8] W. Rianingsih, M. Hasan & A. Pradjaningsih, Portfolio Optimization Using Quadratic Programming, (Jurnal Majalah Ilmiah Matematika dan Statistika, Indonesia, 2017) pp. 67 – 78
- [9] Nurwahidah, Quadratic Programming: An Optimization Tool For Building Global Minimum Variance Portfolio With No Short Sale, (Barengkeng, Indonesia, 2021), pp. 305 – 314
- [10] J.H. Mustakini. Portfolio Theory and Investment Analysis, (BPFE, Indonesia, 2014)

Homomorphic Ordering of The Max-plus Formal Polynomials and The Max-plus Polynomial Functions

Urai Fiqri Fahturrahman⁽¹⁾, Nilamsari Kusumastuti^{*(1)}, and Bayu Prihandono⁽¹⁾

⁽¹⁾Mathematics Department, Tanjungpura University, Indonesia

*Corresponding author: nilamsari@math.untan.ac.id

Abstract. Max-plus algebra, denoted by \mathbb{R}_{\max} , is a set $\mathbb{R} \cup \{-\infty\}$ with two binary operations, maximum and addition, represented as \oplus and \otimes , respectively. Several studies have shown that max-plus algebra is a semifield. The set of n -dimensional vectors over \mathbb{R}_{\max} , denoted by $(\mathbb{R}_{\max})^n$ is a moduloid over \mathbb{R}_{\max} . Moduloid is an analogy of vector space in conventional algebra. Max-plus algebra has a different structure than classic algebra, making further study interesting. The concepts in \mathbb{R}_{\max} can be adapted to formal polynomials and polynomial functions over max-plus algebra. Therefore, this study aims to analyze the structure and the ordered relations on the set of formal polynomials over max-plus algebra ($\mathbb{R}_{\max}[\gamma]$) and the set of polynomial functions over max-plus algebra ($\mathcal{P}(\mathbb{R}_{\max})$). The starting point is to analyze the structure of $\mathbb{R}_{\max}[\gamma]$ and $\mathcal{P}(\mathbb{R}_{\max})$. Then, we define the binary relation \leq_1 on $\mathbb{R}_{\max}[\gamma]$ and the binary relation \leq_2 on $\mathcal{P}(\mathbb{R}_{\max})$, which holds the partial ordered relation axioms. Furthermore, we analyze the ordered relation between $(\mathbb{R}_{\max}[\gamma], \leq_1)$ and $(\mathcal{P}(\mathbb{R}_{\max}), \leq_2)$. This research shows that $(\mathbb{R}_{\max}[\gamma], \leq_1)$ and $(\mathcal{P}(\mathbb{R}_{\max}), \leq_2)$ are partially ordered idempotent algebra (POIAs), and these two POIAs are orderly homomorphic.

Keywords: partial order, idempotent algebra, moduloid, POIA, orderly homomorphic.

INTRODUCTION

Mathematics is one of the disciplines that is fundamental for the development of sciences. In mathematics, there are various studies, one of them is algebra. Algebra is a mathematics study that focuses on the application of symbols or letters in simplifying and solving problems. One of the topics in algebra is abstract algebra, which studies algebraic structures. The application of abstract algebra can be extended to several studies. De Schutter and Van Den Boom explained that, there has been more interest in the industry and the academic world in techniques to model, analyze and control discrete event systems (DES). The technique is production systems, railway networks, traffic control networks and queuing systems [1], [2].

Furthermore, although DES produces a nonlinear description in conventional algebra, the model becomes linear when formulated in max-plus algebra. Max-plus algebra was introduced by Baccelli, Cohen, Olsder and Quadrat as written in a book entitled Synchronization and Linearity. Max-plus algebra uses maximum and addition as its binary operations. Specifically, DES can be modelled by using the maximum operation, in which a new event starts as soon as all preceding operations have been finished and the addition operation, in which an event's finishing time equals the event's starting time plus the duration [3]. This leads to the linear description in the max-plus algebra, referred to as the max-plus linear DES.

Analyzing the max-plus linear DES model can be done by using the properties of max-plus algebra. Max-plus algebra, denoted by \mathbb{R}_{\max} is the set $\mathbb{R} \cup \{-\infty\}$, denoted by \mathbb{R}_{ε} with the maximum operation as \oplus and the addition operation as \otimes [4]. Max-plus algebra is a commutative idempotent semifield where the zero element is $-\infty$ denoted by ε , and the unity element is 0 as e . The concept of max-plus algebra can be extended to another concept. The max-plus linear system has been studied [5]–[7]. The matrix theory [8]–[11] and graph representation in max-plus algebra have been developed. Meanwhile, the polynomial theory in max-plus algebra has also been studied [12].

This paper will consider the homomorphic ordering of the max-plus formal polynomials and the max-plus polynomial functions. Based on the materials in the max-plus algebra, partial orders on the idempotent algebra of formal polynomials and polynomial functions are introduced to derive two partially ordered idempotent algebras (POIAs). The orderly structural relationships among these two POIAs are then studied. It is proven that the POIA of formal polynomials and the POIA of polynomial functions are orderly homomorphic.

MATERIALS AND METHODS

There is some relation in binary relation, and one of which is a partial order. A binary relation \leq is called a partial order on S if \leq is reflexive, antisymmetric and transitive [13]. A set of S with partial order \leq is called partially ordered set (poset). If $(S, +)$ is idempotent commutative semigroup, then the binary relation \leq on S where $x, y \in S$, $x \leq y \Leftrightarrow x + y = y$ is a partial order on S [14], [15]. Before discussing about idempotent algebra, let us see the definition of moduloid.

Definition 1 [12] Let a commutative monoid (M, \oplus_1) with zero element ε_1 and idempotent semifield (K, \oplus_2, \otimes_2) with zero element ε_2 and unity element e_2 . A set of M is called a moduloid over K if there exists a mapping

$$\begin{aligned} \varphi & : K \times M \rightarrow M \\ & : (k, a) \mapsto k \otimes a \end{aligned}$$

Satisfying these properties:

- i) $\varphi(k, a \oplus_1 b) = \varphi(k, a) \oplus_1 \varphi(k, b)$,
- ii) $\varphi(k \oplus_2 l, a) = \varphi(k, a) \oplus_1 \varphi(l, a)$,
- iii) $\varphi(k, \varphi(l, a)) = \varphi(k \otimes_2 l, a)$,
- iv) $\varphi(e_2, a) = a$,
- v) $\varphi(\varepsilon_2, a) = \varepsilon_1$,

where $a, b \in M$ and $k, l \in K$.

Definition 1 leads us to an idempotent algebra explained in Definition 2.

Definition 2 [12] Let a moduloid (M, \oplus_1) over idempotent semifield (K, \oplus_2, \otimes_2) . A moduloid (M, \oplus_1) with internal operation \otimes_1 is called idempotent algebra if

- i) \otimes_1 is associative.
- ii) \otimes_1 is distributive with respect to \oplus_1 .
- iii) (M, \otimes_1) has an identity element e_1 .

Now, we discuss about partially ordered idempotent algebra.

Definition 3 [12] Let an idempotent algebra (M, \oplus, \otimes) over semifield K . (M, \oplus, \otimes) with a partial order \leq is called Partially Ordered Idempotent Algebra (POIA) if $a_1 \leq b_1$ and $a_2 \leq b_2$, then

- i) $a_1 \oplus a_2 \leq b_1 \oplus b_2$,
- ii) $a_1 \otimes a_2 \leq b_1 \otimes b_2$,
- iii) $ka_1 \leq kb_1$,

where $a_1, a_2, b_1, b_2 \in M$ and $k \in K$.

Furthermore, according to Definition 3, we obtained that (M, \oplus, \otimes) is a POIA. It then defined a homomorphic mapping between two idempotent algebras. Detail is explained in Definition 4.

Definition 4 [12] Let $(M_1, \oplus_1, \otimes_1)$ and $(M_2, \oplus_2, \otimes_2)$ are two idempotent algebras over semifield K . A mapping $\varphi : M_1 \rightarrow M_2$ is a homomorphic if

- i) $\varphi(a \oplus_1 b) = \varphi(a) \oplus_2 \varphi(b)$,
- ii) $\varphi(a \otimes_1 b) = \varphi(a) \otimes_2 \varphi(b)$,
- iii) $\varphi(ka) = k\varphi(a)$,

where $a, b \in M_1$ and $k \in K$.

Next, we analyze the ordered between two POIAs as in Definition 5.

Definition 5 [12] Two POIAs, (M_1, \leq_1) and (M_2, \leq_2) , are orderly homomorphic if there exists a homomorphism σ from M_1 to M_2 such that the following propositions hold:

- i) For every $a, b \in M_1$, if $a \leq_1 b$, then $\sigma(a) \leq_2 \sigma(b)$,
- ii) For every $u, v \in M_2$, if $u \leq_2 v$, then there exist $a, b \in M_1$ such that $\sigma(a) = u$, $\sigma(b) = v$ and $a \leq_1 b$.

According to Definition 5, then two POIAs, (M_1, \leq_1) and (M_2, \leq_2) , are orderly homomorphic. Max-plus algebra is a set of $\mathbb{R} \cup \{-\infty\}$ with \oplus and \otimes as binary operation. Max-plus algebra, denoted as $\mathbb{R}_{\max} = (\mathbb{R}_\varepsilon, \oplus, \otimes)$ where $\mathbb{R}_\varepsilon = \mathbb{R} \cup \{-\infty\}$, \oplus is maximum operation and \otimes is addition operation, defined as $a \oplus b = \max\{a, b\}$ and $a \otimes b = a + b$, where $a, b \in \mathbb{R}_\varepsilon$ [16]. The order of operations is quite similar to in classical algebra, which \otimes is prioritized then \oplus . Max-plus algebra is commutative idempotent semifield with $-\infty$ as zero element denoted by ε and 0 as unity element denoted by e [17]. The power of multiplication and division operation in \mathbb{R}_{\max} defined as

$$x^{\otimes n} = \underbrace{x \otimes x \otimes \dots \otimes x}_{n \text{ times}} \text{ and } \frac{x}{y} = x \otimes y^{\otimes -1} = x - y,$$

respectively.

RESULTS AND DISCUSSION

The max-plus formal polynomial is a formal polynomial over max-plus algebra, a set of infinite sequences

$$p = (p_0, p_1, p_2, \dots) \quad (1)$$

where $p_l \in \mathbb{R}_\varepsilon$, $l \in \mathbb{N} \cup \{0\}$, and there exists $n \in \mathbb{N} \cup \{0\}$ such $p_l = \varepsilon$ for $l > n$, for \mathbb{N} is the set of natural number. For every $p = (p_0, p_1, p_2, \dots)$ and $q = (q_0, q_1, q_2, \dots)$, the internal operation defined by:

- i) $p \oplus q = (p_0, p_1, p_2, \dots) \oplus (q_0, q_1, q_2, \dots) = (p_0 \oplus q_0, p_1 \oplus q_1, p_2 \oplus q_2, \dots)$
- ii) $p \otimes q = (p_0, p_1, p_2, \dots) \otimes (q_0, q_1, q_2, \dots) = (r_0, r_1, r_2, \dots)$ with $r_l = \bigoplus_{i+j=l} p_i \otimes q_j$.

Basically, $d \in \mathbb{R}_\varepsilon$ can be represented as $(d, \varepsilon, \varepsilon, \dots)$. Then, based on ii), the external operation is defined by:

- iii) $d \otimes p = (d, \varepsilon, \varepsilon, \dots) \otimes (p_0, p_1, p_2, \dots) = (d \otimes p_0, d \otimes p_1, d \otimes p_2, \dots)$.

Next, according to Lemma 6, (1) can be represented as $p(\gamma) = \bigoplus_{l=0}^n p_l \otimes \gamma^{\otimes l}$.

Lemma 6 Let a formal polynomial $\gamma = (\gamma_0, \gamma_1, \gamma_2, \dots)$ where

$$\gamma_l = \begin{cases} e, & l = 1 \\ \varepsilon, & l \neq 1 \end{cases}$$

then, for every integers $n \geq 1$, formal polynomial $\gamma^{\otimes n}$ where

$$(\gamma^{\otimes n})_l = \begin{cases} e, & l = n \\ \varepsilon, & l \neq n \end{cases}$$

holds true.

Proof. We will prove by induction as follows

- i) For $n = 1$, then

$$(\gamma^{\otimes 1})_l = \gamma_l = \begin{cases} e, & l = 1 \\ \varepsilon, & l \neq 1 \end{cases} \quad (2)$$

Thus, it is true for $n = 1$.

- ii) Assume true for $n = k$ such that

$$(\gamma^{\otimes k})_l = \begin{cases} e, & l = k \\ \varepsilon, & l \neq k \end{cases} \quad (3)$$

then for $n = k + 1$,

$$(\gamma^{\otimes(k+1)})_l = (\gamma^{\otimes k} \otimes \gamma)_l = \bigoplus_{i+j=l} (\gamma^{\otimes k})_i \otimes \gamma_j$$

where $l \in \mathbb{N} \cup \{0\}$. Based on (2) and (3), then for $l = k + 1$,

$$\begin{aligned} (\gamma^{\otimes(k+1)})_{k+1} &= \bigoplus_{i+j=k+1} (\gamma^{\otimes k})_i \otimes \gamma_j \\ &= (\gamma^{\otimes k})_0 \otimes \gamma_{k+1} \oplus (\gamma^{\otimes k})_1 \otimes \gamma_k \oplus \dots \oplus (\gamma^{\otimes k})_k \otimes \gamma_1 \oplus \\ &\quad (\gamma^{\otimes k})_{k+1} \otimes \gamma_0 \\ &= \varepsilon \oplus \varepsilon \oplus \dots \oplus e \oplus \varepsilon \\ &= e. \end{aligned}$$

When for $l \neq k + 1$, then

$$\begin{aligned} (\gamma^{\otimes(k+1)})_l &= \bigoplus_{i+j=l} (\gamma^{\otimes k})_i \otimes \gamma_j \\ &= (\gamma^{\otimes k})_0 \otimes \gamma_l \oplus (\gamma^{\otimes k})_1 \otimes \gamma_{l-1} \oplus \dots \oplus (\gamma^{\otimes k})_{l-1} \otimes \gamma_1 \oplus \\ &\quad (\gamma^{\otimes k})_l \otimes \gamma_0 \\ &= \varepsilon \oplus \varepsilon \oplus \dots \oplus (\gamma^{\otimes k})_{l-1} \otimes \gamma_1 \oplus \varepsilon \\ &= (\gamma^{\otimes k})_{l-1} \otimes \gamma_1 \\ &= (\gamma^{\otimes k})_{l-1} \otimes e \\ &= (\gamma^{\otimes k})_{l-1}. \end{aligned}$$

Since $l \neq k + 1$, then $l - 1 \neq k$. Thus, based on (3), $(\gamma^{\otimes(k+1)})_l = (\gamma^{\otimes k})_{l-1} = \varepsilon$, for every $l \neq k + 1$.

Hence, it is true for $n = k + 1$, so

$$(\gamma^{\otimes(k+1)})_l = \begin{cases} e, & l = k + 1 \\ \varepsilon, & l \neq k + 1 \end{cases} \quad (4)$$

Therefore, according to i) and ii),

$$(\gamma^{\otimes n})_l = \begin{cases} e, & l = n \\ \varepsilon, & l \neq n \end{cases}$$

holds true for every integers $n \geq 1$. ■

In particular, as usual convention $\gamma^{\otimes n}$ with $n = 0$ can be seen as $\gamma^{\otimes 0} = (e, \varepsilon, \varepsilon, \varepsilon, \dots)$ and identified as constant e . Thus, formal polynomial p can be represented into

$$\begin{aligned} (p_0, p_1, \dots, p_n, \varepsilon, \dots) &= (p_0, \varepsilon, \varepsilon, \dots) \oplus (\varepsilon, p_1, \varepsilon, \dots) \oplus \dots \oplus (\varepsilon, \dots, p_n, \varepsilon, \dots) \\ &= p_0 \otimes (e, \varepsilon, \varepsilon, \dots) \oplus p_1 \otimes (\varepsilon, e, \varepsilon, \dots) \oplus \dots \oplus p_n \otimes (\varepsilon, \dots, e, \varepsilon, \dots) \\ &= p_0 \otimes \gamma^{\otimes 0} \oplus p_1 \otimes \gamma^{\otimes 1} \oplus \dots \oplus p_n \otimes \gamma^{\otimes n} \\ &= \bigoplus_{l=0}^n p_l \otimes \gamma^{\otimes l} = p(\gamma) \end{aligned} \quad (5)$$

The set of max-plus formal polynomials denoted by $\mathbb{R}_{\max}[\gamma]$, defined as follows:

$$\mathbb{R}_{\max}[\gamma] = \{p(\gamma) \mid p(\gamma) = \bigoplus_{l=0}^n p_l \otimes \gamma^{\otimes l}\},$$

where zero element is $\varepsilon(\gamma) \in \mathbb{R}_{\max}[\gamma]$, $\varepsilon_l = \varepsilon$ for $l \in \mathbb{N} \cup \{0\}$, and unity element is $e(\gamma) \in \mathbb{R}_{\max}[\gamma]$, $e_l = \begin{cases} e, & l = 0 \\ \varepsilon, & l \neq 0 \end{cases}$ for $l \in \mathbb{N} \cup \{0\}$. For every $p(\gamma) = \bigoplus_{l=0}^n p_l \otimes \gamma^{\otimes l}$, $q(\gamma) = \bigoplus_{l=0}^m q_l \otimes \gamma^{\otimes l} \in \mathbb{R}_{\max}[\gamma]$ where $n \geq m \in \mathbb{N} \cup \{0\}$, $d \in \mathbb{R}_{\max}$, the internal and external operations of $\mathbb{R}_{\max}[\gamma]$ are defined as follows:

a) Internal operations:

i) Addition: $p(\gamma) \oplus q(\gamma) = \bigoplus_{l=0}^n (p_l \oplus q_l) \otimes \gamma^{\otimes l}$

ii) Multiplication: $p(\gamma) \otimes q(\gamma) = \bigoplus_{l=0}^{n+m} (\bigoplus_{i+j=l} p_i \otimes q_j) \otimes \gamma^{\otimes l}$

b) External operation: $d \otimes p(\gamma) = \bigoplus_{l=0}^n (d \otimes p_l) \otimes \gamma^{\otimes l}$.

Theorem 7 The set of max-plus formal polynomials is idempotent algebra.

Proof. Let

$$\begin{aligned} \varphi &: \mathbb{R}_{\max} \times \mathbb{R}_{\max}[\gamma] \rightarrow \mathbb{R}_{\max}[\gamma] \\ &: (a, p(\gamma)) \mapsto a \otimes p(\gamma) \end{aligned}$$

If $(a, p(\gamma)) = (b, q(\gamma))$, then $\varphi((a, p(\gamma))) = a \otimes p(\gamma) = b \otimes q(\gamma) = \varphi((b, q(\gamma)))$, thus φ well defined.

Next, for every $p(\gamma), q(\gamma) \in \mathbb{R}_{\max}[\gamma]$ and $a \in \mathbb{R}_{\max}$ then

$$\begin{aligned} \text{i) } \varphi(a, p(\gamma) \oplus q(\gamma)) &= a \otimes (p(\gamma) \oplus q(\gamma)) \\ &= (a \otimes p(\gamma)) \oplus (a \otimes q(\gamma)) \\ &= \varphi(a, p(\gamma)) \oplus \varphi(a, q(\gamma)), \end{aligned}$$

$$\begin{aligned} \text{ii) } \varphi(a \oplus b, p(\gamma)) &= (a \oplus b) \otimes p(\gamma) \\ &= (a \otimes p(\gamma)) \oplus (b \otimes p(\gamma)) \\ &= \varphi(a, p(\gamma)) \oplus \varphi(b, p(\gamma)), \end{aligned}$$

$$\begin{aligned} \text{iii) } \varphi(a, \varphi(b, p(\gamma))) &= a \otimes (b \otimes p(\gamma)) \\ &= (a \otimes b) \otimes p(\gamma) \\ &= \varphi(a \otimes b, p(\gamma)), \end{aligned}$$

$$\begin{aligned} \text{iv) } \varphi(e, p(\gamma)) &= e \otimes p(\gamma) \\ &= p(\gamma), \end{aligned}$$

$$\begin{aligned} \text{v) } \varphi(\varepsilon, p(\gamma)) &= \varepsilon \otimes p(\gamma) \\ &= \varepsilon(\gamma). \end{aligned}$$

So, $(\mathbb{R}_{\max}[\gamma], \oplus)$ is a moduloid over \mathbb{R}_{\max} . Next, we define \otimes as new internal operation on $(\mathbb{R}_{\max}[\gamma], \oplus)$, such for every $p(\gamma), q(\gamma), r(\gamma) \in \mathbb{R}_{\max}[\gamma]$, then

$$\begin{aligned} \text{i) } (p(\gamma) \otimes q(\gamma)) \otimes r(\gamma) &= p(\gamma) \otimes q(\gamma) \otimes r(\gamma) \\ &= p(\gamma) \otimes (q(\gamma) \otimes r(\gamma)) \end{aligned}$$

ii) Right-distributive holds:

$$p(\gamma) \otimes (q(\gamma) \oplus r(\gamma)) = (p(\gamma) \otimes q(\gamma)) \oplus (p(\gamma) \otimes r(\gamma))$$

and left-distributive also holds:

$$(p(\gamma) \oplus q(\gamma)) \otimes r(\gamma) = (p(\gamma) \otimes r(\gamma)) \oplus (q(\gamma) \otimes r(\gamma))$$

iii) Let $e(\gamma)$ as unity element in $\mathbb{R}_{\max}[\gamma]$. It is easy to show that $e(\gamma)$ where $e_l = \begin{cases} e, & l = 0 \\ \varepsilon, & l \neq 0 \end{cases}$ such

$e(\gamma) \otimes p(\gamma) = p(\gamma) \otimes e(\gamma) = p(\gamma)$. Thus, $\mathbb{R}_{\max}[\gamma]$ has a unity element.

Hence, $(\mathbb{R}_{\max}[\gamma], \oplus, \otimes)$ is idempotent algebra over \mathbb{R}_{\max} . ■

Formal polynomial (2) is associated with a polynomial function as follows:

$$\hat{p} : \mathbb{R}_{\max} \rightarrow \mathbb{R}_{\max}$$

$$: c \mapsto \hat{p}(c) = \bigoplus_{l=0}^n p_l \otimes c^{\otimes l}.$$

If $c = d \in \mathbb{R}_{\max}$, then $\hat{p}(c) = \bigoplus_{l=0}^n p_l \otimes c^{\otimes l} = \bigoplus_{l=0}^n p_l \otimes d^{\otimes l} = \hat{p}(d)$, thus \hat{p} well defined. The set of polynomial functions over \mathbb{R}_{\max} denoted by $\mathcal{P}(\mathbb{R}_{\max})$, defined as follows:

$$\mathcal{P}(\mathbb{R}_{\max}) = \{\hat{p}: \mathbb{R}_{\max} \rightarrow \mathbb{R}_{\max} \mid \hat{p}(c) = \bigoplus_{l=0}^n p_l \otimes c^{\otimes l}, \forall c \in \mathbb{R}_{\max}\}$$

where zero element is $\hat{\varepsilon} \in \mathcal{P}(\mathbb{R}_{\max})$, $\hat{\varepsilon}(c) = \bigoplus_{l=0}^n \varepsilon_l \otimes c^{\otimes l}$ and unity element is $\hat{e} \in \mathcal{P}(\mathbb{R}_{\max})$, $\hat{e}(c) = \bigoplus_{l=0}^n e_l \otimes c^{\otimes l}$.

For every $\hat{p}, \hat{q} \in \mathcal{P}(\mathbb{R}_{\max})$ and $c, d \in \mathbb{R}_{\max}$, the internal and external operations of $\mathcal{P}(\mathbb{R}_{\max})$ are defined as follows:

a) Internal operations:

$$\text{i) Addition: } (\hat{p} \oplus \hat{q})(c) = \hat{p}(c) \oplus \hat{q}(c)$$

$$\text{ii) Multiplication: } (\hat{p} \otimes \hat{q})(c) = \hat{p}(c) \otimes \hat{q}(c)$$

b) External operations: $(d \otimes \hat{p})(c) = d \otimes \hat{p}(c)$.

Theorem 8 The set of max-plus polynomial functions is idempotent algebra.

Proof. Let

$$\varphi : \mathbb{R}_{\max} \times \mathcal{P}(\mathbb{R}_{\max}) \rightarrow \mathcal{P}(\mathbb{R}_{\max})$$

$$: (a, \hat{p}) \mapsto a \otimes \hat{p}$$

If $(a, \hat{p}) = (b, \hat{q})$, then $\varphi((a, \hat{p})) = a \otimes \hat{p} = b \otimes \hat{q} = \varphi((b, \hat{q}))$, thus φ well defined. Next, for every $\hat{p}, \hat{q} \in \mathcal{P}(\mathbb{R}_{\max})$ and $a, c \in \mathbb{R}_{\max}$, then

$$\begin{aligned} \text{i) } \varphi(a, \hat{p} \oplus \hat{q})(c) &= a \otimes (\hat{p} \oplus \hat{q})(c) \\ &= a \otimes (\hat{p}(c) \oplus \hat{q}(c)) \\ &= (a \otimes \hat{p}(c)) \oplus (a \otimes \hat{q}(c)) \\ &= \varphi(a, \hat{p})(c) \oplus \varphi(a, \hat{q})(c) \end{aligned}$$

Since for all $c \in \mathbb{R}_{\max}$. Thus, $\varphi(a, \hat{p} \oplus \hat{q}) = \varphi(a, \hat{p}) \oplus \varphi(a, \hat{q})$,

$$\begin{aligned} \text{ii) } \varphi(a \oplus b, \hat{p})(c) &= (a \oplus b) \otimes \hat{p}(c) \\ &= (a \otimes \hat{p}(c)) \oplus (b \otimes \hat{p}(c)) \\ &= \varphi(a, \hat{p})(c) \oplus \varphi(b, \hat{p})(c) \end{aligned}$$

Since for all $c \in \mathbb{R}_{\max}$. Thus, $\varphi(a \oplus b, \hat{p}) = \varphi(a, \hat{p}) \oplus \varphi(b, \hat{p})$,

$$\begin{aligned} \text{iii) } \varphi(a, \varphi(b, \hat{p})(c))(c) &= a \otimes (b \otimes \hat{p}(c))(c) \\ &= a \otimes (b \otimes \hat{p}(c)) \\ &= (a \otimes b) \otimes \hat{p}(c) \\ &= \varphi(a \otimes b, \hat{p})(c) \end{aligned}$$

Since for all $c \in \mathbb{R}_{\max}$. Thus, $\varphi(a, \varphi(b, \hat{p})) = \varphi(a \otimes b, \hat{p})$,

$$\begin{aligned} \text{iv) } \varphi(e, \hat{p})(c) &= e \otimes \hat{p}(c) \\ &= \hat{p}(c) \end{aligned}$$

Since for all $c \in \mathbb{R}_{\max}$. Thus, $\varphi(e, \hat{p}) = \hat{p}$,

$$\begin{aligned} \text{v) } \varphi(\varepsilon, \hat{p})(c) &= \varepsilon \otimes \hat{p}(c) \\ &= \hat{\varepsilon}(c) \end{aligned}$$

Since for all $c \in \mathbb{R}_{\max}$. Thus, $\varphi(\varepsilon, \hat{p}) = \hat{\varepsilon}$.

So, $(\mathcal{P}(\mathbb{R}_{\max}), \oplus)$ is a moduloid over \mathbb{R}_{\max} . Next, we define \otimes as new internal operation on $(\mathcal{P}(\mathbb{R}_{\max}), \oplus)$, such for every $\hat{p}, \hat{q}, \hat{r} \in \mathcal{P}(\mathbb{R}_{\max})$ and $c \in \mathbb{R}_{\max}$ then

$$\begin{aligned} \text{i) } (\hat{p}(c) \otimes \hat{q}(c)) \otimes \hat{r}(c) &= \hat{p}(c) \otimes \hat{q}(c) \otimes \hat{r}(c) \\ &= \hat{p}(c) \otimes (\hat{q}(c) \otimes \hat{r}(c)) \end{aligned}$$

Since for all $c \in \mathbb{R}_{\max}$. Thus, $(\hat{p} \otimes \hat{q}) \otimes \hat{r} = \hat{p} \otimes (\hat{q} \otimes \hat{r})$.

ii) Right-distributive holds:

$$\hat{p}(c) \otimes (\hat{q}(c) \oplus \hat{r}(c)) = (\hat{p}(c) \otimes \hat{q}(c)) \oplus (\hat{p}(c) \otimes \hat{r}(c))$$

Since for all $c \in \mathbb{R}_{\max}$. Thus, $\hat{p} \otimes (\hat{q} \oplus \hat{r}) = (\hat{p} \otimes \hat{q}) \oplus (\hat{p} \otimes \hat{r})$. and left-distributive also holds:

$$(\hat{p}(c) \oplus \hat{q}(c)) \otimes \hat{r}(c) = (\hat{p}(c) \otimes \hat{r}(c)) \oplus (\hat{q}(c) \otimes \hat{r}(c))$$

Since for all $c \in \mathbb{R}_{\max}$. Thus, $(\hat{p} \oplus \hat{q}) \otimes \hat{r} = (\hat{p} \otimes \hat{r}) \oplus (\hat{q} \otimes \hat{r})$.

iii) Let \hat{e} as unity element in $\mathcal{P}(\mathbb{R}_{\max})$. It is easy to show that $\hat{e}(c)$, where $e_l = \begin{cases} e, & l = 0 \\ \varepsilon, & l \neq 0 \end{cases}$ such $\hat{e}(c) \otimes \hat{p}(c) = \hat{p}(c) \otimes \hat{e}(c) = \hat{p}(c)$. Since, for all $c \in \mathbb{R}_{\max}$. Thus, $\mathcal{P}(\mathbb{R}_{\max})$ has a unity element.

Hence, $(\mathcal{P}(\mathbb{R}_{\max}), \oplus, \otimes)$ is idempotent algebra over \mathbb{R}_{\max} . ■

Next, we define a binary relation \preceq_1 on $\mathbb{R}_{\max}[\gamma]$ as follows:

for every $p(\gamma), q(\gamma) \in \mathbb{R}_{\max}[\gamma]$, $p(\gamma) \preceq_1 q(\gamma)$ if and only if $p_l \leq q_l$, for $l \in \mathbb{N} \cup \{0\}$.

Lemma 9 The binary relation \preceq_1 is partial order on $\mathbb{R}_{\max}[\gamma]$.

Proof. $p_l = p_l$ for $l \in \mathbb{N} \cup \{0\}$, obviously. Thus, $p(\gamma) \preceq_1 p(\gamma)$, so \preceq_1 is reflexive. If $p(\gamma) \preceq_1 q(\gamma)$ then $p_l \leq q_l$, and if $q(\gamma) \preceq_1 p(\gamma)$, then $q_l \leq p_l$. Thus, $p_l = q_l$ for $l \in \mathbb{N} \cup \{0\}$, so $p(\gamma) = q(\gamma)$. Hence, \preceq_1 is antisymmetric. If $p(\gamma) \preceq_1 q(\gamma)$ then $p_l \leq q_l$, and if $q(\gamma) \preceq_1 r(\gamma)$ then $q_l \leq r_l$. Thus, $p_l \leq r_l$ for $l \in \mathbb{N} \cup \{0\}$, so $p(\gamma) \preceq_1 r(\gamma)$. Hence, \preceq_1 is transitive. Therefore, \preceq_1 is a partial order. ■

Lemma 10 For every $p(\gamma), q(\gamma) \in \mathbb{R}_{\max}[\gamma]$ and partial order \leq_1 , $p(\gamma) \leq_1 q(\gamma)$ if and only if $p(\gamma) \oplus q(\gamma) = q(\gamma)$.

Proof. (\Rightarrow) Since $p(\gamma) \leq_1 q(\gamma)$, then $p_l \leq q_l$ for all $l \in \mathbb{N} \cup \{0\}$. Thus, $\max\{p_l, q_l\} = q_l$ if and only if $p_l \oplus q_l = q_l$. Then,

$$p(\gamma) \oplus q(\gamma) = \bigoplus_{l=0}^n (p_l \oplus q_l) \otimes \gamma^{\otimes l} = \bigoplus_{l=0}^m q_l \otimes \gamma^{\otimes l} = q(\gamma).$$

(\Leftarrow) Since $p(\gamma) \oplus q(\gamma) = q(\gamma)$, then $p_l \oplus q_l = q_l$. Thus, $p_l \leq q_l$ for all $l \in \mathbb{N} \cup \{0\}$, so $p(\gamma) \leq_1 q(\gamma)$. Since (\Rightarrow) dan (\Leftarrow) has been proved, then $p(\gamma) \leq_1 q(\gamma)$ if and only if $p(\gamma) \oplus q(\gamma) = q(\gamma)$. ■

Theorem 11 $(\mathbb{R}_{\max}[\gamma], \leq_1)$ is a Partially Ordered Idempotent Algebra (POIA).

Proof. Based on Lemma 10, for all $p(\gamma), p'(\gamma), q(\gamma), q'(\gamma) \in \mathbb{R}_{\max}[\gamma]$, if $p(\gamma) \leq_1 q(\gamma)$ and $p'(\gamma) \leq_1 q'(\gamma)$, then $p(\gamma) \oplus q(\gamma) = q(\gamma)$ and $p'(\gamma) \oplus q'(\gamma) = q'(\gamma)$, respectively. Thus,

i) $(p(\gamma) \oplus p'(\gamma)) \oplus (q(\gamma) \oplus q'(\gamma)) = (p(\gamma) \oplus q(\gamma)) \oplus (p'(\gamma) \oplus q'(\gamma)) = q(\gamma) \oplus q'(\gamma)$,
then $p(\gamma) \oplus p'(\gamma) \leq_1 q(\gamma) \oplus q'(\gamma)$. For $i, j \in \mathbb{N} \cup \{0\}$, then

$$\begin{aligned} q_i \otimes q'_j &= (p_i \oplus q_i) \otimes (p'_j \oplus q'_j) \\ &= p_i \otimes p'_j \oplus p_i \otimes q'_j \oplus (q_i \otimes p'_j \oplus q_i \otimes q'_j) \\ &= p_i \otimes p'_j \oplus p_i \otimes q'_j \oplus (q_i \otimes (p'_j \oplus q'_j)) \\ &= p_i \otimes p'_j \oplus (p_i \otimes q'_j \oplus q_i \otimes q'_j) \\ &= p_i \otimes p'_j \oplus ((p_i \oplus q_i) \otimes q'_j) \\ &= p_i \otimes p'_j \oplus q_i \otimes q'_j \end{aligned}$$

Thus, for $l \in \mathbb{N} \cup \{0\}$,

$$\begin{aligned} ((p \otimes p') \oplus (q \otimes q'))_l &= (p \otimes p')_l \oplus (q \otimes q')_l \\ &= (\bigoplus_{i+j=l} p_i \otimes p'_j) \oplus (\bigoplus_{i+j=l} q_i \otimes q'_j) \\ &= \bigoplus_{i+j=l} (p_i \otimes p'_j \oplus q_i \otimes q'_j) \\ &= \bigoplus_{i+j=l} (q_i \otimes q'_j) \\ &= (q \otimes q')_l \end{aligned}$$

hence, we obtain

ii) $(p(\gamma) \otimes p'(\gamma)) \oplus (q(\gamma) \otimes q'(\gamma)) = q(\gamma) \otimes q'(\gamma)$,

then $p(\gamma) \otimes p'(\gamma) \leq_1 q(\gamma) \otimes q'(\gamma)$. For $l \in \mathbb{N} \cup \{0\}$, $d \in \mathbb{R}_{\max}$, then

$$\begin{aligned} (d \otimes p)_l \oplus (d \otimes q)_l &= d \otimes p_l \oplus d \otimes q_l \\ &= d \otimes (p_l \oplus q_l) \\ &= d \otimes q_l \end{aligned}$$

hence, we obtain

iii) $(d \otimes p(\gamma)) \oplus (d \otimes q(\gamma)) = d \otimes q(\gamma)$,

then, $d \otimes p(\gamma) \leq_1 d \otimes q(\gamma)$. Therefore, $(\mathbb{R}_{\max}[\gamma], \leq_1)$ is Partially Ordered Idempotent Algebra (POIA). ■

Next, we define a binary relation \leq_2 on $\mathcal{P}(\mathbb{R}_{\max})$ as follows:

for every $\hat{p}, \hat{q} \in \mathcal{P}(\mathbb{R}_{\max})$, $\hat{p} \leq_2 \hat{q}$ if and only if $\hat{p}(c) \leq \hat{q}(c)$, $c \in \mathbb{R}_{\max}$.

Lemma 12 The binary relation \leq_2 is a partial order on $\mathcal{P}(\mathbb{R}_{\max})$.

Proof. $\hat{p}(c) = \hat{p}(c)$ for $c \in \mathbb{R}_{\max}$, obviously. Thus, $\hat{p} \leq_2 \hat{p}$, so \leq_2 is reflexive. If $\hat{p} \leq_2 \hat{q}$ then $\hat{p}(c) \leq \hat{q}(c)$, and if $\hat{q} \leq_2 \hat{p}$, then $\hat{q}(c) \leq \hat{p}(c)$. Thus, $\hat{p}(c) = \hat{q}(c)$ for $c \in \mathbb{R}_{\max}$, then $\hat{p} = \hat{q}$. Hence, \leq_2 is antisymmetric. If $\hat{p} \leq_2 \hat{q}$ then $\hat{p}(c) \leq \hat{q}(c)$, and if $\hat{q} \leq_2 \hat{r}$ then $\hat{q}(c) \leq \hat{r}(c)$. Thus, $\hat{p}(c) \leq \hat{r}(c)$ for $c \in \mathbb{R}_{\max}$, so $\hat{p} \leq_2 \hat{r}$. Hence, \leq_2 is transitive. Therefore, \leq_2 is a partial order. ■

Theorem 13 $(\mathcal{P}(\mathbb{R}_{\max}), \leq_2)$ is a Partially Ordered Idempotent Algebra (POIA).

Proof. For $\hat{p}, \hat{q} \in \mathcal{P}(\mathbb{R}_{\max})$ and $c \in \mathbb{R}_{\max}$, if $\hat{p}(c) \leq \hat{q}(c)$, then $\hat{p}(c) \oplus \hat{q}(c) = \hat{q}(c)$. Since for all $c \in \mathbb{R}_{\max}$, then $\hat{p} \oplus \hat{q} = \hat{q}$. Thus, based on Lemma 12, for $\hat{p}_1, \hat{p}_2, \hat{q}_1, \hat{q}_2 \in \mathcal{P}(\mathbb{R}_{\max})$, if $\hat{p}_1 \leq_2 \hat{q}_1$ and $\hat{p}_2 \leq_2 \hat{q}_2$, then $\hat{p}_1 \oplus \hat{q}_1 = \hat{q}_1$ and $\hat{p}_2 \oplus \hat{q}_2 = \hat{q}_2$, respectively. Thus,

i) $(\hat{p}_1 \oplus \hat{p}_2) \oplus (\hat{q}_1 \oplus \hat{q}_2) = (\hat{p}_1 \oplus \hat{q}_1) \oplus (\hat{p}_2 \oplus \hat{q}_2) = \hat{q}_1 \oplus \hat{q}_2$

then, $\hat{p}_1 \oplus \hat{p}_2 \leq_2 \hat{q}_1 \oplus \hat{q}_2$. Next,

ii) $(\hat{p}_1 \otimes \hat{p}_2) \oplus (\hat{q}_1 \otimes \hat{q}_2) = \hat{p}_1 \otimes \hat{p}_2 \oplus ((\hat{p}_1 \oplus \hat{q}_1) \otimes (\hat{p}_2 \oplus \hat{q}_2))$

$$\begin{aligned}
 &= \hat{p}_1 \otimes \hat{p}_2 \oplus (\hat{p}_1 \otimes \hat{p}_2 \oplus \hat{p}_1 \otimes \hat{q}_2 \oplus \hat{q}_1 \otimes \hat{p}_2 \oplus \hat{q}_1 \otimes \hat{q}_2) \\
 &= (\hat{p}_1 \otimes \hat{p}_2 \oplus \hat{p}_1 \otimes \hat{p}_2) \oplus \hat{p}_1 \otimes \hat{q}_2 \oplus \hat{q}_1 \otimes \hat{p}_2 \oplus \hat{q}_1 \otimes \hat{q}_2 \\
 &= (\hat{p}_1 \otimes \hat{p}_2 \oplus \hat{p}_1 \otimes \hat{q}_2) \oplus (\hat{q}_1 \otimes \hat{p}_2 \oplus \hat{q}_1 \otimes \hat{q}_2) \\
 &= \hat{p}_1 \otimes (\hat{p}_2 \oplus \hat{q}_2) \oplus \hat{q}_1 \otimes (\hat{p}_2 \oplus \hat{q}_2) \\
 &= \hat{p}_1 \otimes \hat{q}_2 \oplus \hat{q}_1 \otimes \hat{q}_2 \\
 &= (\hat{p}_1 \oplus \hat{q}_1) \otimes \hat{q}_2 \\
 &= \hat{q}_1 \otimes \hat{q}_2
 \end{aligned}$$

then $\hat{p}_1 \otimes \hat{p}_2 \preceq_2 \hat{q}_1 \otimes \hat{q}_2$. Next,

$$\begin{aligned}
 \text{iii) } (d \otimes \hat{p}_1) \oplus (d \otimes \hat{q}_1) &= d \otimes (\hat{p}_1 \oplus \hat{q}_1) \\
 &= d \otimes \hat{q}_1
 \end{aligned}$$

then $d \otimes \hat{p}_1 \preceq_2 d \otimes \hat{q}_1$. Therefore, $(\mathcal{P}(\mathbb{R}_{\max}), \preceq_2)$ is Partially Ordered Idempotent Algebra (POIA). ■

Theorem 14 Two POIAs $(\mathbb{R}_{\max}[\gamma], \preceq_1)$ and $(\mathcal{P}(\mathbb{R}_{\max}), \preceq_2)$ are orderly homomorphic.

Proof. Let σ be a function as follows:

$$\begin{aligned}
 \sigma &: \mathbb{R}_{\max}[\gamma] \rightarrow \mathcal{P}(\mathbb{R}_{\max}) \\
 &: p(\gamma) \mapsto \sigma(p(\gamma)) = \hat{p}.
 \end{aligned}$$

If $p(\gamma) = q(\gamma)$, then $\sigma(p(\gamma)) = \hat{p} = \hat{q} = \sigma(q(\gamma))$, thus σ well defined. Next, for every $p(\gamma), q(\gamma) \in \mathbb{R}_{\max}[\gamma]$ and $c, d \in \mathbb{R}_{\max}$, then

$$\begin{aligned}
 \text{i) } \sigma(p(\gamma) \oplus q(\gamma))(c) &= (\widehat{p \oplus q})(c) \\
 &= \bigoplus_{l=0}^n (p \oplus q)_l \otimes c^{\otimes l} \\
 &= \bigoplus_{l=0}^n (p_l \oplus q_l) \otimes c^{\otimes l} \\
 &= \left(\bigoplus_{l=0}^n p_l \otimes c^{\otimes l} \right) \oplus \left(\bigoplus_{l=0}^m q_l \otimes c^{\otimes l} \right) \\
 &= \hat{p}(c) \oplus \hat{q}(c) \\
 &= \sigma(p(\gamma))(c) \oplus \sigma(q(\gamma))(c)
 \end{aligned}$$

Since for all $c \in \mathbb{R}_{\max}$. Thus, $\sigma(p(\gamma) \oplus q(\gamma)) = \sigma(p(\gamma)) \oplus \sigma(q(\gamma))$

$$\begin{aligned}
 \text{ii) } \sigma(p(\gamma) \otimes q(\gamma))(c) &= (\widehat{p \otimes q})(c) \\
 &= \bigoplus_{l=0}^{n+m} (p \otimes q)_l \otimes c^{\otimes l} \\
 &= \bigoplus_{l=0}^{n+m} \left(\bigoplus_{i+j=l} p_i \otimes q_j \right) \otimes c^{\otimes l} \\
 &= \left(\bigoplus_{l=0}^n p_l \otimes c^{\otimes l} \right) \otimes \left(\bigoplus_{l=0}^m q_l \otimes c^{\otimes l} \right) \\
 &= \hat{p}(c) \otimes \hat{q}(c) \\
 &= \sigma(p(\gamma))(c) \otimes \sigma(q(\gamma))(c)
 \end{aligned}$$

Since for all $c \in \mathbb{R}_{\max}$. Thus, $\sigma(p(\gamma) \otimes q(\gamma)) = \sigma(p(\gamma)) \otimes \sigma(q(\gamma))$

$$\begin{aligned}
 \text{iii) } \sigma(d \otimes p(\gamma))(c) &= (\widehat{d \otimes p})(c) \\
 &= \bigoplus_{l=0}^n (d \otimes p)_l \otimes c^{\otimes l} \\
 &= \bigoplus_{l=0}^n d \otimes p_l \otimes c^{\otimes l} \\
 &= d \otimes \bigoplus_{l=0}^n p_l \otimes c^{\otimes l}
 \end{aligned}$$

$$\begin{aligned} &= d \otimes \hat{p}(c) \\ &= d \otimes \sigma(p(\gamma))(c) \end{aligned}$$

Since for all $c \in \mathbb{R}_{\max}$. Thus, $\sigma(d \otimes p(\gamma)) = d \otimes \sigma(p(\gamma))$.

Hence, σ is homomorphism. Let $p(\gamma), q(\gamma) \in \mathbb{R}_{\max}[\gamma]$, since $p(\gamma) \preceq_1 q(\gamma)$, then $p_l \leq q_l \Leftrightarrow p_l \oplus q_l = q_l$. Thus,

$$\begin{aligned} \sigma(p(\gamma))(c) \oplus \sigma(q(\gamma))(c) &= \hat{p}(c) \oplus \hat{q}(c) \\ &= \left(\bigoplus_{l=0}^n p_l \otimes c^{\otimes l} \right) \oplus \left(\bigoplus_{l=0}^m q_l \otimes c^{\otimes l} \right) \\ &= \bigoplus_{l=0}^m (p_l \oplus q_l) \otimes c^{\otimes l} \\ &= \bigoplus_{l=0}^m q_l \otimes c^{\otimes l} \\ &= \hat{q}(c) \\ &= \sigma(q(\gamma))(c) \end{aligned}$$

Since $\sigma(p(\gamma))(c) \oplus \sigma(q(\gamma))(c) = \sigma(q(\gamma))(c)$, then $\sigma(p(\gamma))(c) \leq \sigma(q(\gamma))(c)$ for all $c \in \mathbb{R}_{\max}$. Therefore, $\sigma(p(\gamma)) \preceq_2 \sigma(q(\gamma))$. Let $\hat{p}, \hat{q} \in \mathcal{P}(\mathbb{R}_{\max})$ and $\hat{p} \preceq_2 \hat{q}$, then there exist $a(\gamma), b(\gamma) \in \mathbb{R}_{\max}[\gamma]$ such that for all $c \in \mathbb{R}_{\max}$ $\sigma(a(\gamma))(c) = \hat{p}(c) = \bigoplus_{l=0}^n p_l \otimes c^{\otimes l}$ and $\sigma(b(\gamma))(c) = \hat{q}(c) = \bigoplus_{l=0}^m q_l \otimes c^{\otimes l}$. It is easy to show that $a(\gamma) = p(\gamma)$ and $b(\gamma) = q(\gamma)$, since p_l and $q_l, l \in \mathbb{N} \cup \{0\}$ are coefficients in $\hat{p}(c)$ and $\hat{q}(c)$, respectively. Since $\hat{p} \preceq_2 \hat{q}$, then $\hat{p}(c) \leq \hat{q}(c)$. Thus,

$$\hat{p}(c) \oplus \hat{q}(c) = \hat{q}(c) \Leftrightarrow \bigoplus_{l=0}^n (p_l \oplus q_l) \otimes c^{\otimes l} = \bigoplus_{l=0}^m q_l \otimes c^{\otimes l}.$$

Hence, $p_l \oplus q_l = q_l$, then $p_l \leq q_l$ such that $a(\gamma) \preceq_1 b(\gamma)$. Therefore, two POIAs $(\mathbb{R}_{\max}[\gamma], \preceq_1)$ and $(\mathcal{P}(\mathbb{R}_{\max}), \preceq_2)$ are orderly homomorphic. ■

CONCLUSION

This paper presents two POIAs of polynomials over \mathbb{R}_{\max} based on properties of binary relations and analyzes the homomorphic ordering among them. The internal and external operations on max-plus formal polynomials and max-plus polynomial functions have been defined. The partial order on them has also been proved. Thus, the max-plus formal polynomials and the max-plus polynomial functions are POIAs. The relationship among them has been defined, which is a homomorphism mapping. It is proven that the POIA of max-plus formal polynomials and the POIA of max-plus polynomial functions are orderly homomorphic.

ACKNOWLEDGEMENT

We would like to thank the Faculty of Mathematics and Natural Science, Tanjungpura University for funding this research. We also thank anonymous referees for several helpful comments and corrections.

REFERENCES

- [1] J. Komenda, S. Lahaye, J.-L. Boimond, and T. Van Den Boom, "Max-plus Algebra in The History of Discrete Event Systems," *Annual Reviews in Control*, vol. 45, pp. 240–249, 2018.
- [2] J. Komenda, S. Lahaye, J.-L. Boimond, and T. Van Den Boom, "Max-Plus Algebra and Discrete Event Systems," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 1784–1790, Jul. 2017.
- [3] R. M. F. Cândido, L. Hardouin, M. Lhommeau, and R. Santos Mendes, "Conditional Reachability of Uncertain Max Plus Linear Systems," *Automatica*, vol. 94, pp. 426–435, Aug. 2018.
- [4] V. M. Gonçalves, C. A. Maia, and L. Hardouin, "On Max-plus Linear Dynamical System Theory: The Observation Problem," *Automatica*, vol. 107, pp. 103–111, Sep. 2019.
- [5] J. Xu, T. Van Den Boom, and B. De Schutter, "Model Predictive Control for Stochastic Max-Plus Linear Systems with Chance Constraints," *IEEE Trans. Automat. Contr.*, vol. 64, no. 1, pp. 337–342, Jan. 2019.
- [6] Y. Tao and C. Wang, "Global Optimization for Max-plus Linear Systems and Applications in Distributed Systems," *Automatica*, vol. 119, p. 109104, Sep. 2020.

- [7] W. Leela-apiradee, P. Thipwiwatpotjana, and A. Gorka, "Closed Form of L-localized Solution Set of Max-plus Interval Linear System and Its Application on Optimization Problem," *Journal of Computational and Applied Mathematics*, vol. 317, pp. 113–127, Jun. 2017.
- [8] S. Taradipa, N. Kusumastuti, and Helmi, "Perbandingan Struktur Himpunan Matriks Interval dan Himpunan Matriks Interval atas Aljabar Max-Plus," *Bimaster*, vol. 11, no. 2, pp. 337–346, 2022.
- [9] P. Butkovič and D. Jones, "On Special Cases of The Generalized Max-Plus Eigenproblem," *SIAM J. Matrix Anal. & Appl.*, vol. 37, no. 3, pp. 1002–1021, Jan. 2016.
- [10] H. Myšková, "Interval Max-plus Matrix Equations," *Linear Algebra and its Applications*, vol. 492, pp. 111–127, Mar. 2016.
- [11] C. Wang and Y. Tao, "Matrix Representation of Formal Polynomials over Max-plus Algebra," *J. Algebra Appl.*, vol. 20, no. 11, p. 2150216, Nov. 2021.
- [12] C. Wang, Y. Xia, and Y. Tao, "Ordered Structures of Polynomials over Max-Plus Algebra," *Symmetry*, vol. 13, no. 7, p. 1137, Jun. 2021.
- [13] A. Wohlgemuth, *Introduction to Proof in Abstract Mathematics*. Newburyport: Dover Publications, 2014.
- [14] M. A. Rudhito, *Aljabar Max-plus dan Penerapannya*. Yogyakarta: Sanata Dharma University Press, 2016.
- [15] S. Kališnik and D. Lešnik, "Symmetric Polynomials in Tropical Algebra Semirings," *Journal of Symbolic Computation*, vol. 93, pp. 100–119, Jul. 2019.
- [16] A. Rosenmann, F. Lehner, and A. Peperko, "Polynomial Convolutions in Max-plus Algebra," *Linear Algebra and its Applications*, vol. 578, pp. 370–401, Oct. 2019.
- [17] C. Wang and Y. Tao, "Interval Strong Solutions of Interval Systems of Max-plus Linear Equations," *Linear Algebra and its Applications*, vol. 537, pp. 148–159, Jan. 2018.

Clustering Analysis of Lecturer Profiles Using Gower Distance and K-Medoids Algorithm

Widya Kusuma Arum¹, Arum Handini Primandari^{*1}

¹Statistics Department, Universitas Islam Indonesia, Yogyakarta, 55584, Indonesia

^{*}Corresponding author: primandari.arum@uii.ac.id

Abstract. A way to preserve the quality of higher education is to improve the quality of human resources through lecturers. Profiling the lectures will help related parties provide suitable policies or programs. In order to do lecturers' profiling, this research used 8,404 lecturers' profiles data obtained from the Higher Education Service Institute (Lembaga Layanan Pendidikan Tinggi abbreviated as LLDIKTI) Region V Yogyakarta, Indonesia. We employed the Gower distance with the K-Medoids procedure. The Gower distance is a metric that measures the dissimilarity of two items with mixed numeric and categorical data. Determining the initial number of clusters operated the elbow methods, resulting in $k=6$. Cluster evaluation measures Dunn Index, Average Silhouette Width, and Within Between Ratio. Lecturer members in each cluster have different characteristics. Lecturers in cluster 1 have the best quality compared to other clusters. Meanwhile, cluster 3 and 4 is a group of lecturers who do not yet have functional positions and lecturer profession certification. Thus lecturers in this cluster require more programs to encourage them.

Keywords: Clustering Analysis, Gower Distance, K-Medoids, Lectures Profiling, Pam

INTRODUCTION

The Indonesian government has made various efforts to develop the education sector at various levels, starting from primary, secondary, and tertiary education. One of the levels that is the focus of attention is higher education because one can face increasingly fierce competition and global competition through higher education. With quality higher education, it is hoped that generations will be born with high abilities, skills, and knowledge, to advance a nation in the future.

Higher education provides higher education after senior secondary education or high school. Higher education is the provider of higher education, which includes diploma, undergraduate, master's, specialist, and doctoral education programs [1]. There are several tertiary institutions in Indonesia, namely Polytechnics, High Schools, Academies, Universities, and institutes. Each type of form usually has many majors or study programs and many lecturers teaching in each of these departments or study programs.

One of the elements in the administration of higher education is the lecturer. The main task of lecturers is as professional educators and scientists responsible for teaching, developing, and disseminating science, technology, and art through education, research, and community service. Lecturers have a position as professionals who play an essential role in enhancing the dignity and development of science, technology, and the arts to improve the quality of national education, implement the national education system and achieve predetermined educational goals [2].

In the Special Province of Yogyakarta, there are 99 private universities. The large number of private tertiary institutions in the Special Province of Yogyakarta also affects the large number of lecturers in the Special Province of Yogyakarta. Based on data from lecturers at the *Lembaga Layanan Pendidikan Tinggi* Region V with the working area covering the Special Region of Yogyakarta Province, the number of private university lecturers in the Special Region of Yogyakarta Province is 8404 lecturers [3].

The large number of lecturers at private tertiary institutions in the Special Province of Yogyakarta certainly makes the characteristics of lecturers more diverse. Each lecturer at each private tertiary institution in the Province of the Special Region of Yogyakarta has different characteristics starting from their functional position, certification status, and the work area where the lecturer teaches. The collection of lecturer datasets can be explored further to see hidden patterns and characteristics and has never been further analyzed (data taken as of February 15, 2023).

Basically, there are two types of grouping, hierarchical and non-hierarchical clustering. Hierarchical Clustering: Hierarchical clustering is basically an unsupervised clustering technique that involves creating clusters in a predefined order. Non Hierarchical Clustering: Non Hierarchical Clustering involves forming new clusters by merging or splitting clusters [4]. The data in this study is categorical data, so the K-medoid non-hierarchical method with a gower distance will be used.

The purpose of this study was to determine the characteristics of private university lecturers in the Special Region of Yogyakarta Province by profiling or grouping lecturers at Private Higher Education in the Special Region of Yogyakarta Province, based on data available at the *Lembaga Layanan Pendidikan Tinggi* (LLDIKTI) Region V by applying the k-medoid clustering method and by using the Gower distance calculation. From the results of this study, the characteristics of the lecturers can be identified so that they can be used as a basis for coaching policies, assistance for lecturers, and encouraging lecturers to be able to develop their careers and competencies immediately.

MATERIALS AND METHODS

Data and Data Sources

The data used in this study is from *Lembaga Layanan Pendidikan Tinggi* Region V. The data itself is academic profile of private university lecturers in DIY Province, which is secondary data. It was obtained through the EvirA LLDIKTI Region V Yogyakarta website, which was taken on 15 February 2023.

Table 1 Definition of Research Variables

Variable	Data Type	Description
Functional Position	Ordinal	The functional position of a lecturer describes a position that shows the duties, responsibilities, authorities, and rights of a lecturer in a higher education unit. Lecturer functional position variables consist of Lecturer, Education Staff, Teaching Staff, Assistants Professor (2), Assistant Professor (1), Associate Professor, and Professor.
Certification Status	Nominal	Certification status is the certification status of lecturers in private tertiary institutions. The certification status variable consists of Not Yet Serdos and Already Serdos.
Region	Nominal	The region explains the location of private tertiary institutions where lecturers teach. Regional variables include Bantul Regency, Gunungkidul Regency, Kulon Progo Regency, Sleman Regency, and Yogyakarta City.

Cluster Analysis

Cluster analysis is a method in multivariate analysis to classify objects based on their characteristics. There are two types of clustering methods for grouping data: hierarchical and non-hierarchical. In the hierarchical clustering method, objects are grouped based on a similarity measure starting from two or more objects that are most similar. Then, other objects with a second affinity will be grouped, forming a tree-like level. In the non-hierarchical clustering method, initial cluster initiation and the following process are carried out without following the hierarchical process [5].

Gower Distance

Gower distance calculation takes the average of the partisan differences between individuals. For each partial dissimilarity in the Gower distance, it has a range between 0 and 1. If the comparison of the results is close to 0, then the two data are identical, or the distance between observations is close together. Conversely, if the comparison is close to 1, the two data are different, or the distance between observations is far apart [6]. The following is the formula for calculating the Gower distance [7]:

$$s_j(x_1, x_2) = 1 - \frac{|y_{1j} - y_{2j}|}{R_j} \quad (1)$$

Description:

- $s_j(x_1, x_2)$ = the distance between the 1st and 2nd data variables
 y_{1j} = the 1st data value on the variable
 y_{2j} = the 2nd data value on the variable
 R_j = variable ranges

Elbow Method

The Elbow method is a method for determining the optimal number of clusters by looking at the percentage comparison between the number of clusters that produces an angled angle at a point [8]. If the difference between the first and second cluster values results in the most significant corner on the graph or decreases, then the number of clusters is considered appropriate. In the elbow method, the Sum of Square Error (SSE) value reduction can also be used to determine the best cluster. The following is the calculation of the SSE value [9]:

$$SSE = \sum_{K=1}^K \sum_{x_i \in S_k} |x_i - C_k|^2 \quad (2)$$

Description:

- K = number of clusters
 x_i = data i
 C_k = centroid cluster

K-Medoids Clustering Algorithm

The K-Medoids algorithm, also called the PAM (Partitioning Around Medoid) algorithm, was created by Leonard Kaufman and Peter J. It is a partitioning clustering technique employed to categorize items into clusters. In this approach, a single item is chosen as the medoid, representing a cluster within the set of items. Cluster formation is achieved by evaluating the proximity between the medoid and non-medoid items [10]. The difference between the K-Means and K-Medoids algorithms relates to how the cluster center is determined. In the K-Means algorithm, the center of each cluster is determined by calculating the average value (mean) of the data points within the cluster. On the other hand, in the K-Medoids algorithm, the center of the cluster, known as the medoid, is identified by selecting a specific data object from the cluster [11].

The K-Medoids algorithm is also used to overcome the weakness of the k-means algorithm, which is sensitive to outlier data. This is because the presence of outliers in the cluster can interfere with the average value (mean) of the cluster because of its distance and characteristics that are different from the majority of other data, so it can distort the average value of the cluster. The K-Medoids algorithm can be done with the following steps [12]:

- Start by initializing k cluster centers (number of clusters).
- Assign each object to the nearest cluster using the Gower distance for distance calculation.
- Once the distances are calculated, randomly select a new cluster center from each object as a non-medoid candidate.
- Calculate the distance of each object in each cluster with the non-medoid candidates.
- Determine the total deviation (S) by calculating the difference between the new and old total distances. If S is less than 0, exchange objects with non-medoid cluster data to create a new set of objects as medoids.
- Repeat steps c-e until there are no further changes in the medoids, resulting in obtaining the clusters and their respective cluster members.

Dunn Index

One of the cluster validation techniques is the Dunn Index. It offers the best score for clustering methods that create clusters with high similarity within a cluster but low similarity between clusters. The ratio of the smallest intra-cluster distance to the largest inter-cluster distance is calculated using this method. Therefore, the number of clusters formed is ideal for a higher Dunn Index value. The Dunn Index is calculated using the following formula [13]:

$$D = \left\{ \frac{\min_{\substack{1 \leq i \leq k \\ i+1 \leq j \leq q}} (d(C_i, C_j))}{\max_{1 \leq l \leq q} d(C_l)} \right\} \quad (3)$$

Description:

$d(C_i, C_j)$: The closeness measure between cluster i and cluster j .

$d(C_l)$: The closeness measure between members in a cluster l .

Average Silhouette Width

Average Silhouette Width (ASW) is a metric used to evaluate the quality of clusterization. It measures the extent to which each object in the cluster is within its group compared to the rest. ASW is calculated for each object in the dataset and then averaged to get an overall value. ASW values range from -1 to 1, with higher values indicating better clusterization quality. The ASW calculation process involves two main components: coherence within and separation between clusters. Within-cluster coherence measures the extent to which objects in the same cluster are close to each other. Inter-cluster separation measures the extent to which objects from different clusters are separated from each other [14].

Within-Between Ratio

In clusterization validation, the within-between ratio is used as one of the metrics to measure the compactness of groups and the distance between groups. This ratio provides information on how well the groups are separated from each other. Generally, the higher the within-between ratio, the better the group separation. This indicates that the within-group variation is much smaller than the between-group variation, making the groups more compact and separated. The calculation of the within-between ratio is:

$$D = \frac{\text{Mean of Suares Between}}{\text{Mean of Suares Within}} \quad (4)$$

Research Flow

The following is the flow of research conducted by researchers.

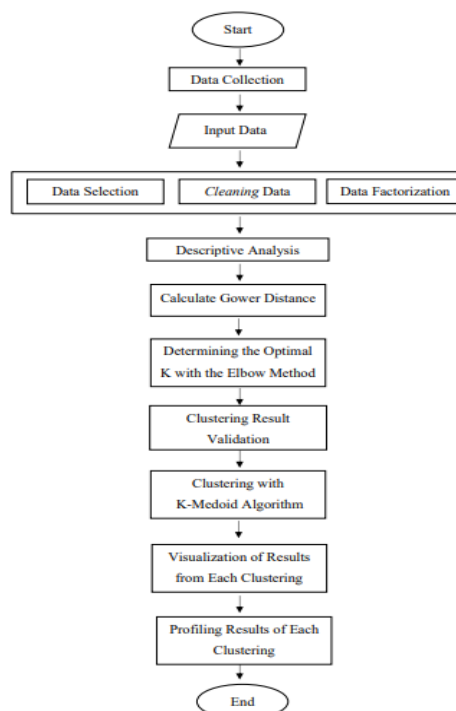


Figure 1. Functional Position of Lecturer

This research was conducted using R-Studio software. The explanation of the stages based on Figure 2 is as follows:

1. Collecting data obtained through the EVirA website and inputting the data used, namely data on lecturers of private tertiary institutions in the Special Province of Yogyakarta.
2. Preprocess data with data selection, cleaning, and factorization. Data with a character form in research is converted into factor form, so it can then be used for K-Medoids clustering analysis.
3. Descriptive analysis of data variables of functional position, certification status, and college area where lecturers teach.
4. Performing distance calculations using Gower distance. Gower distance measures the similarity or difference between objects in the data.
5. Determining K-Optimal with the elbow method. In the Elbow method, the best cluster value is taken from the Sum of Square Error (SSE) value, which has decreased significantly and is elbow-shaped.
6. Perform validation of clusters to determine the optimal cluster formed using the Dunn Index, Average Silhouette Width, and Within-Between Ratio.
7. Perform clustering with K-Medoid.

At this stage, K-Medoid clustering testing uses the cluster and factoextra libraries contained in the RStudio software.

8. Visualizing the clustering results to make it easier to read and analyze the characteristics of each cluster obtained.
9. Profiling the results of each cluster formed.
The cluster results are profiled to identify and compare the characteristics of lecturers in each cluster formed.

RESULTS AND DISCUSSION

Descriptive Analysis

Descriptive analysis is used to determine the general description and information of the variables used in the research. The following graph illustrates the functional position variables, lecturer certification status, and private college area where lecturers teach at private tertiary institutions in the Special Province of Daerah Istimewa Yogyakarta.

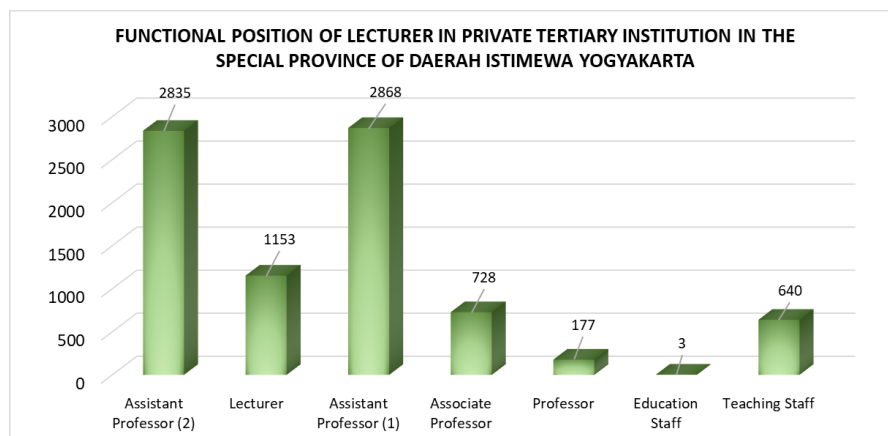


Figure 2. Functional Position of Lecturer

Based on Figure 1, it can be seen that lecturers with functional positions as assistant professor (1) have the highest number, namely 34.13% of the total private tertiary institution lecturers in the Province of the Special Region of Yogyakarta. Lecturers with functional positions as assistant professor (2) also have a high number after an assistant professor (1), namely 33.73%. Meanwhile, lecturers with functional positions as professors have a significant difference compared to assistant professor (1) and assistant professor (2), namely only 2.13% of the total lecturers. This shows that only a few private university lecturers in the Province of the Special Region of Yogyakarta have achieved the highest position. Lecturers with functional positions as educational staff have the least number compared to other functional positions, namely only 0.04% of the total number of lecturers at private tertiary institutions in the Province of the Special Region of Yogyakarta.

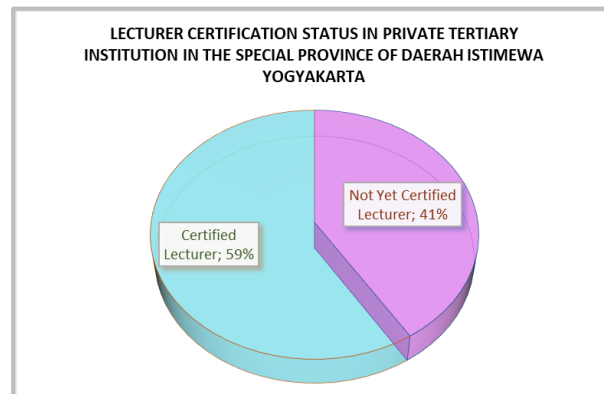


Figure 3. Lecturer Certification Status

Based on the circle diagram of the percentage of lecturer certification status in Figure 2, it can be seen that the percentage of lecturers who already have lecturer certification status is 59% of the total lecturers at private tertiary institutions in the Province of DIY, which is higher when compared to lecturers who do not yet have lecturer certification. Lecturers who do not yet have lecturer certification account for 41% of the total private tertiary institution lecturers in the Province of the Special Region of Yogyakarta.

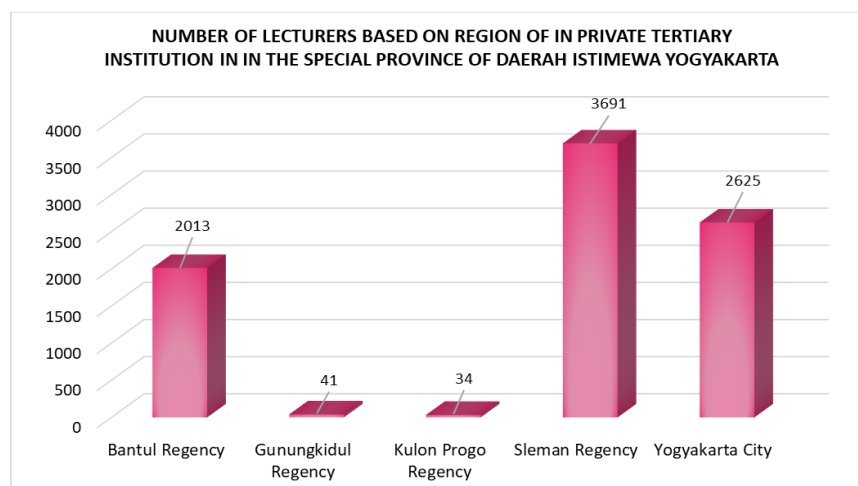


Figure 4. Lecturer Based on Region

Based on Figure 3, it can be seen that the percentage of lecturers teaching in private tertiary institutions in Sleman Regency is high compared to other provinces, namely 43.92% of the total lecturers in private tertiary institutions in the Special Province of Yogyakarta. The city of Yogyakarta has a quite significant percentage, namely 31.24% of the total lecturers. Meanwhile, lecturers who teach in the Gunung Kidul and Kulon Progo Regencies have a low percentage compared to other districts, namely only 0.49% and 0.40%. This happened because the number of private tertiary institution in the two districts was small.

Gower Distance

Gower distance is used to measure the similarity of data objects in k-medoid clustering. The concept in Gower distance is relatively simple because a specific distance metric that works well for each variable type is used and scaled between 0 and 1. Then, linear combinations using the simplest user-defined weights are averaged to create the final distance matrix. The following is the table results of the Gower distance used in this study.

Table 2 Lecturer Grouping Distribution

	1	2	3	4	5	...	8403	8404
1	0.0000000	0.0000000	0.0000000	0.3333333	0.3333333	...	0.6666667	0.6666667
2	0.0000000	0.0000000	0.0000000	0.3333333	0.3333333	...	0.6666667	0.6666667
3	0.0000000	0.0000000	0.0000000	0.3333333	0.3333333	...	0.6666667	0.6666667
4	0.3333333	0.3333333	0.3333333	0.0000000	0.3333333	...	0.6666667	0.6666667
5	0.3333333	0.3333333	0.3333333	0.3333333	0.0000000	...	0.3333333	0.3333333
.
8403	0.6666667	0.6666667	0.6666667	0.6666667	0.3333333	...	0.0000000	0.3333333
8404	0.6666667	0.6666667	0.6666667	0.6666667	0.3333333	...	0.3333333	0.0000000

The distance between the data object and itself is 0. This can be seen from the matrix's main diagonal, where the diagonal values are 0.0000000 with itself. Based on the table, it can also be seen that the distance between data objects 1 and 2 is 0.0000000. This shows that data objects 1 and 2 have a perfect resemblance.

Elbow Method

The Elbow method was used to test the optimal number of clusters formed in this study. From Figure 4, it can be seen that when the number of clusters is $k = 1$, the Total Within SSE (Sum of Square Error) value is the highest compared to the SSE values at the other k . Meanwhile, when the number of clusters $k = 2$, the SSE value decreased significantly, and when the number of clusters $k = 3$, the SSE value again reduced. This occurred until the number of clusters $k = 10$. Based on the graph of the elbow method formed, it can be seen that the number of clusters that form the elbow is visible when the number of clusters is $k = 5$ or when the number of clusters is $k = 6$. Therefore, the optimal number of clusters formed can be between $k = 5$ or when the number of clusters is $k = 6$.

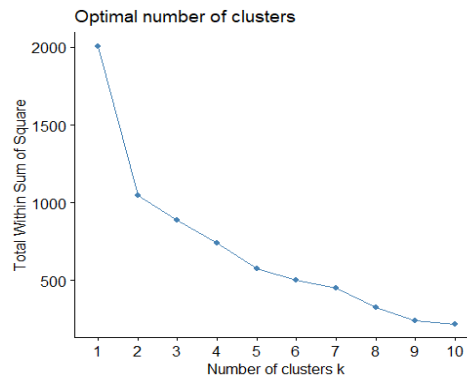


Figure 5. Plot Elbow Method

Cluster Result Validation

To determine the optimal cluster between $k = 5$ and $k = 6$, a comparison of validation values is carried out to get the best cluster results. For comparison of validation values can be seen in Table 3.

Table 3 Lecturer Grouping Distribution

Validation	Cluster	
	K=5	K=6
Dunn Index	0.3333333	0.3333333
Average Silhouette Width	0.4129712	0.4167469
Within-Between Ratio	0.41042	0.3726919

Based on the table above, it can be seen that in the clustering of private university lecturers with the k-medoids method with three validations, the best clustering is obtained, namely:

1. For Dunn index validation, the validation value is the same between k=5 and k=6, namely 0.3333333.
2. For the Average Silhouette Width validation, the maximum validation value is obtained at k=6, which is 0.4167469.
3. To make sure the validation results of the Average Silhouette Width are validated with Within-Between Ratio and the minimum validation value is obtained at k = 6, which is 0.3726919.

From the validation results obtained based on the three validation Dunn Index, Average Silhouette Width, and Within-Between Ratio, it can be concluded that the best clustering using these three validations is clustering at k=6, because it has the most optimum validation value in each validation.

K-Medoid Clustering Results with k=6

After determining the optimal cluster using the elbow method and performing cluster validation, the results of clustering k-medoid k = 6 are described in table 4.

Table 4 Lecturer Grouping Distribution

	Count Of Lecturers	Percentage of Lecturers
Cluster 1	2385	28.38%
Cluster 2	1221	14.53%
Cluster 3	1169	13.91%
Cluster 4	1465	17.43%
Cluster 5	1157	13.77%
Cluster 6	1007	11.98%

Based on the results of the cluster analysis, it can be seen that cluster 1 has the largest number of lecturer members compared to the other clusters, namely 28.38% of the total private university lecturers in the Special Province of Yogyakarta. Meanwhile, cluster 6 has the least number of lecturers compared to the other clusters, namely only 11.98% of the total private university lecturers in the Special Province of Yogyakarta.

Profiling of K-Medoid Clustering Results

Profiling of cluster analysis results is carried out on the best clustering results. At the profiling stage, the characteristics of each cluster formed will be seen so that trends in each cluster can be seen.

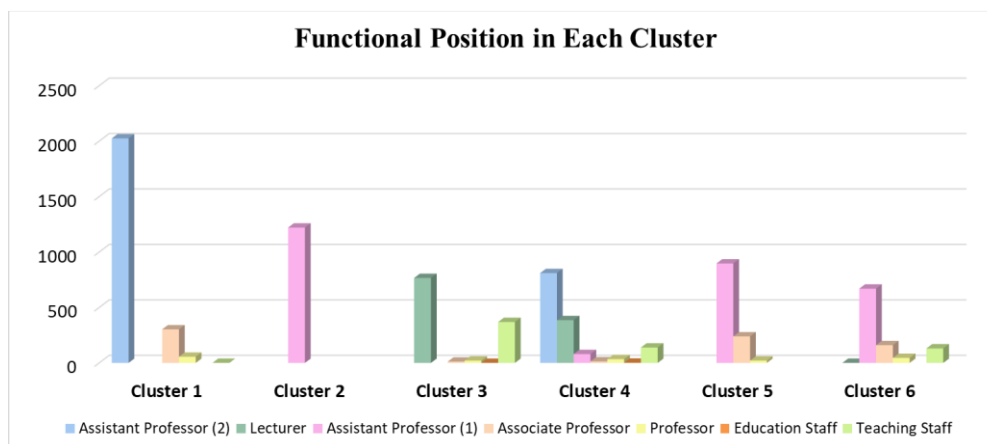


Figure 6. Functional in Each Cluster

From the results of the profiling of the functional positions of the lecturers, there are significant differences between the groups of lecturers in the five clusters formed. In cluster 1, the number of lecturers

with functional positions as assistant professor (2) is the highest compared to other clusters. This shows that in cluster 1, many lecturers are still in the early stages of their careers. In addition, in cluster 1, lecturers with functional positions as associate professor and professor also have many lecturer members compared to other cluster members. This shows that cluster 1 has lecturers with good functional positions compared to other clusters. In clusters 2, 5, and 6, lecturers with the functional position of assistant professor (1) have the most number.

Meanwhile, in clusters 3 and 4, there are still many lecturers who do not have functional positions when compared to other clusters. Because of this, lecturers in clusters 3 and 4 require special attention to apply for functional positions immediately. In cluster 3, the number of lecturers with functional positions as teaching staff has the highest number compared to other clusters. This shows that lecturers in cluster 3 work as teaching staff with contracts or are new lecturers.

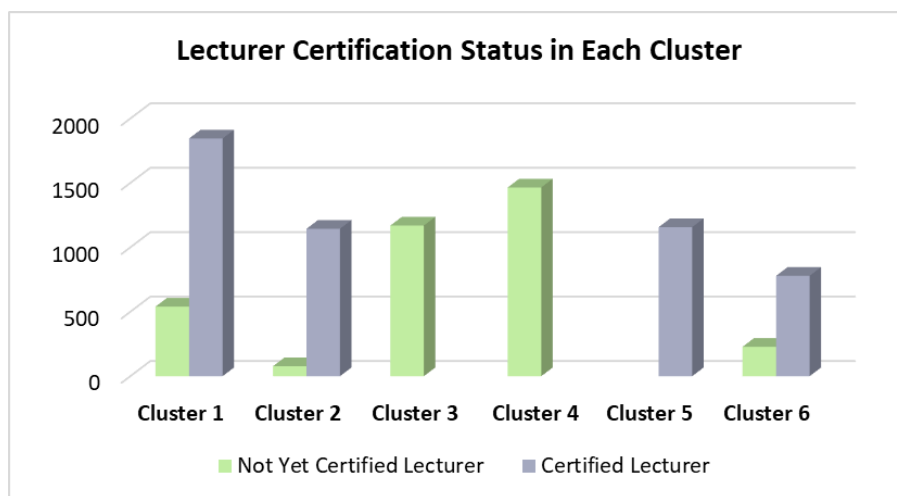


Figure 7. Lecturers Certification Status in Each Cluster

Based on the certification status of PTS lecturers in DIY Province, clusters 1, 2, 5, and 6 have a higher number of lecturers with certification status than those who have not been certified. Meanwhile, in clusters 3 and 4, the number of lecturers who have not been certified has a higher number than lecturers who have been certified. Even in these two clusters, all members have certification status, not yet certified lecturers. From the many lecturers who have yet to be certified in cluster 3 and cluster 4, the lecturers in the members of these two clusters need more encouragement to apply for lecturer certification to support their careers immediately.

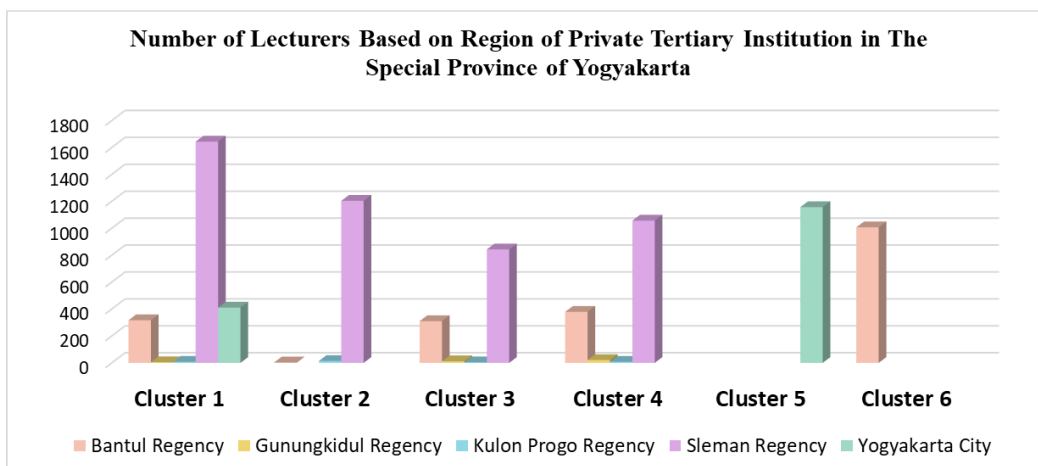


Figure 8. Number of Lecturers Based Region in Each Cluster

Based on the results of profiling the areas where lecturers teach, clusters 1, 2, 3, and 4 have the highest number of lecturers teaching in the Sleman Regency Private Universities area compared to other clusters. The

Sleman Regency area has a large number of private universities, so many lecturers teach in the area. Based on the profiling results, it can also be seen that lecturers in cluster 5 teach the most in Yogyakarta City. The number of lecturers who teach in Gunung Kidul and Kulon Progo regencies has a significant difference in each cluster compared to other categories. This is because the number of private universities in the province is small.

CONCLUSION

From the research that has been done on clustering using k-medoid and gower distance for profiling private higher education lecturers in the Special Province of Yogyakarta, the conclusions are as follows:

1. Lecturers with functional positions as Lectors dominate Private University in DIY Province. Many PTS lecturers in the Special Province of Yogyakarta have already obtained lecturer certification. Based on the area of the private university, most lecturers teach in Sleman Regency.
2. The Elbow method is used to determine the optimal cluster formed in k-medoid clustering. After performing the elbow method and cluster validation using the Dunn index, average silhouette width, and wb ratio, it can be concluded that the best clustering using these three validations is clustering at $k=6$. Based on the results of the cluster analysis, it can be seen that cluster 1 has the most number of lecturers compared to other clusters. Meanwhile, cluster 6 has the least number of lecturers compared to other clusters.
3. Each cluster formed in this study has different lecturer characteristics. Lecturers in cluster 1 have the best members compared to other clusters. This is because in cluster 1, lecturers with functional positions as associate professor and professor have the most number of lecturer members compared to other clusters. Clusters 2, 5, and 6 have better faculty member compared to clusters 3 and 4. Lecturers in clusters 3 and 4 require more special attention than others. This is because clusters 3 and 4 are dominated by lecturers who do not have functional positions, and there are still many lecturers who have not been certified.

ACKNOWLEDGEMENT

We thank to Statistics Laboratory for research facility support and Statistics Department, UII for the research foundation.

REFERENCES

- [1] Pemerintah Indonesia, Undang-Undang Republik Indonesia Nomor 20 Tahun 2003 Tentang Sistem Pendidikan Nasional Pp. 1–7, Jakarta, 2003.
- [2] Peraturan Pemerintah Republik Indonesia Nomor 37, Jakarta: Pemerintah Pusat , 2009.
- [3] "Evira," [Online]. Available: <https://lldikti5.id/evira/frontpage/dosen>. [Accessed 10 Februari 2023].
- [4] F. Hestuningtias and A. H. Primandari, "Pengelompokan Kecamatan Di Kabupaten Merangin Berdasarkan Produksi Tanaman Perkebunan Tahun 2021 Menggunakan Agglomerative Hierarchical Clustering," *Emerging Statistics and Data Science Journa*, vol. 1, no. 1, pp. 137-147, 2023.
- [5] Y. C. and H. Yang, "Research on K-value Selection Method of K-Means CLustering Algorithm," *J* 2.2, pp. 226-235, 2019.
- [6] A. Skabar, "Clustering Mixed-Attribute Data using Random Walk," *International Conference on Computational Science*, pp. 989-997, 2017.
- [7] S. Irsalina, Implementasi Simple K-Medoids Partitioning Pada Pengelompokan Sma Berdasarkan Komponen Kesiapan Teknologi Informasi, Yogyakarta: Universitas Islam Indonesia, 2022.
- [8] N. P. E. Merliana, E. and J. Susanto, "Analisa Penentuan Jumlah Cluster Terbaik Pada Metode K-Means Clustering," *Seminar Nasional Multi Disiplin Ilmu*, pp. 978-979, 2015.
- [9] A. T. Rahman, W. and R. Anggrainingsih, "Coal Trade Data Clustering Using K-Means," *ITSMART: Jurnal Ilmiah Teknologi dan Informasi*, pp. 24-31, 2017.
- [10] L. Kaufman and P. J. Rousseeuw, Finding Groups In Data, New York: John Wiley & Sons, Inc, 1990.

-
- [11] N. K. Kaur and D. Singh, "K-Medoids Clustering Algorithm- A Review," *International Journal of Computer Application and Technology*, vol. I, no. 1, pp. 42-45, 2014.
 - [12] A. Bhat, "K-medoids Clustering Using Partitioning Around Medoids for Performing Face Recognition," *International Journal of Soft Computing Mathematic and Control*, vol. III, no. 3, pp. 1-12, 2014.
 - [13] S. I. Pratiwi, T. Widiharih and A. R. Hakim, "Analisis," *Analisis Klaster Metode Ward dan Average Linkage dengan Validasi Dunn Index dan Koefisien Korelasi Cophenetic (Studi Kasus: Kecelakaan Lalu Lintas Berdasarkan Jenis Kendaraan Tiap Kabupaten/Kota di Jawa tengah Tahun 2018)*, vol. VIII, no. 4, pp. 486-495, 2019.
 - [14] F. Batool and C. Henning, "Clustering with the Average Silhouette Width," *Computational Statistics & Data Analysis*, 2021.

Numerical Simulation of the Tsunami Wave Propagation Model Using Lax-Friedrichs Method

Yulian Fauzi^{*(1)}, Zulfia Memi Mayasari⁽¹⁾, and Agus Susanto⁽²⁾

⁽¹⁾Mathematics Department, FMIPA, University of Bengkulu,
WR. Supratman, Kandang Limun, Bengkulu, 38371 Indonesia

⁽²⁾Math Graduate, FMIPA, University of Bengkulu,
WR. Supratman, Kandang Limun, Bengkulu, 38371 Indonesia

*Corresponding author: yulianfauzi@unib.ac.id

Abstract. Tsunami waves are a series of long waves induced by unexpected and impulsive changes in seawater. This research was conducted to simulate the propagation of tsunami waves using the Lax-Friedrichs method. Tsunami waves are modeled using the shallow water equation (SWE) which consists of the continuity equation and the momentum equation. The differential equation model of tsunami wave propagation consists of wave height, discharge flux height in the x -axis direction, and discharge flux height in the y -axis direction. The step taken is discretize the wave differential equation using the Lax-Friedrichs scheme. The simulation is carried out by input an example case and input the data into the Matlab program that has been made. The amplitude, wavelength, and time will affect the height of the tsunami waves that are formed. Simulation of the tsunami wave propagation model was carried out with an amplitude of 8 meters and a wave length of 50 km. The simulation results show that the height of a tsunami wave is directly proportional to the length of time it takes for the water wave to calm down. The Lax-Friedrichs method is a good method used to predict the behavior of tsunami waves.

Keywords: Discretization, Lax-Friedrichs, Matlab, Simulation, Tsunami Wave Propagation.

INTRODUCTION

Tsunamis are a series of long waves induced by unexpected and impulsive changes in seawater [1]. According to [2], tsunami waves are long waveforms. Long wave propagation can be modeled mathematically using the shallow water wave equation. Flouri et.al [3] states that shallow water waves are waves that arise in shallow water, namely waves whose wavelength is quite large compared to their depth.

The simplest water wave theory that approximates real ocean wave behavior is a system of pairwise partial differential equations known as the shallow water wave equation [4]. Tsunami waves are classified as shallow water waves so their equations are modeled using the shallow water wave equation. Many researchers have modeled the propagation and inundation of tsunami waves. Imamura [5] has conducted studies to obtain a tsunami model that can be completed manually. Various methods have been widely used for modeling the propagation of tsunami waves. Analytical modeling of tsunami wave propagation consists of mathematical equations derived from hydrodynamic theory. This theory is useful for analyzing sea waves when a tsunami occurs. In several concepts about tsunamis [6], [7], [8], [9], [10], it can be formulated that the propagation of tsunami waves from the earthquake source to the mainland depends on the characteristics of the tsunami waves which are influenced by bathymetry, topography, and other factors. obstacles (coastal roughness, seabed material).

The propagation of tsunami waves can be modeled numerically using partial differential equations. The equation used is the shallow water equation (SWE) based on the theory of water wave propagation developed from the theory of fluid mechanics, including the law of conservation of mass and Newton's second law [5], [11]. SWE can be derived from the law of the conservation of mass and the conservation of momentum expressed by the Navier-Stokes equation. Jamhuri [12] in his research derived a two-dimensional shallow water wave equation to simulate how a tsunami wave propagates in all directions generated by an external energy using the law of conservation of mass and the law of conservation of momentum. Fauzi et.al [13] conducted a numerical simulation of the propagation of tsunami waves using the Runge Kutta Geometric Mean (RKGGM) method. Fauzi and Mayasari [14] have investigated the propagation of tsunami waves by discretizing the shallow water wave equation using the Leapfrog method. These three studies use a finite difference method approach in modeling the propagation of tsunami waves.

This research was conducted to simulate the propagation of tsunami waves using the Lax-Friedrichs method. This research is a development of research conducted by [13], [14] using the Lax-friedrichs method through a finite difference approach. Numerical simulations were carried out to see the propagation of tsunami waves using the scenarios carried out by [13].

MATERIALS AND METHODS

Methods

This research begins by examining the basic equations for tsunami waves, the behavior of tsunami waves. This equation relates to the wave height, wave speed, and energy distribution in the domain. Modeling the propagation of tsunami waves using a numerical approach begins with discriminating the tsunami wave equation that derived from the shallow water equation. The Lax-Friedrichs method uses the central difference approach to approximate the spatial derivatives of the basic equations. This involves calculating the difference in the value of a variable between neighboring cells and dividing by the cell size. This approach provides better estimates of the spatial changes in variables. Iterations are carried out at each time step until the desired simulation time is reached. At each iteration step, the variable values are updated based on the calculations performed in the previous step. Calculation results such as wave height at any time and certain location are recorded for further analysis, such as modeling the pattern of the tsunami wave distribution. After the simulation is completed, the results are evaluated and compared with observational data or other simulation results to validate the model. If the simulation results match expectations and observations, the Lax-Friedrichs method can be used to predict the behavior of tsunami waves in different scenarios.

RESULTS AND DISCUSSION

The SWE model forms the characteristics of the sea in several layers that have a constant density value with the topography of the seabed as the lower boundary and the free surface as the upper boundary [5]. The formation of the tsunami propagation model from SWE is based on a physical and mathematical model with the following assumptions:

1. The horizontal scale (wavelength) is larger than the vertical scale (water depth).
2. Viscosity is neglected.
3. Incompressibility (the density of water does not depend on pressure and time).

The consequence of this assumption is that the vertical acceleration of the water particles is negligible compared to the acceleration due to gravity. Consequently the vertical motion of the water particles has no effect on the pressure. The tsunami wave propagation equation [15] is a system of partial differential equations with 3 equations and 3 variables which can be expressed by the following equation:

$$\frac{\partial \eta}{\partial t} + \frac{\partial M}{\partial x} + \frac{\partial N}{\partial y} = 0 \quad (1)$$

$$\frac{\partial M}{\partial t} + \frac{\partial}{\partial x} \left(\frac{M^2}{D} \right) + \frac{\partial}{\partial y} \left(\frac{MN}{D} \right) + gD \frac{\partial \eta}{\partial x} + \frac{gn^2}{D^{7/3}} M \sqrt{M^2 + N^2} = 0 \quad (2)$$

$$\frac{\partial N}{\partial t} + \frac{\partial}{\partial x} \left(\frac{MN}{D} \right) + \frac{\partial}{\partial y} \left(\frac{N^2}{D} \right) + gD \frac{\partial \eta}{\partial y} + \frac{gn^2}{D^{7/3}} N \sqrt{M^2 + N^2} = 0 \quad (3)$$

where x, y : horizontal axis, t : time, η : vertical displacement above the water surface, g : acceleration due to gravity, M : flux of discharge in x direction, N : flux of discharge in y direction, D : total depth of water given by $h + \eta$, n : seabed friction coefficient values.

The equation for the propagation of tsunami waves is analytically very difficult to solve, so the solution is done numerically using the Lax-Friedrichs method. Solving differential equations using the Lax-Friedrichs method can be approached by the finite difference method. To describe the numerical scheme for the tsunami propagation model, namely by discretizing the three-dimensional equation from the shallow water equation using the Lax-Friedrichs method. Procedures in the Lax-Friedrichs method for solving partial differential equations apply the center difference method and the forward difference method. Central or central difference approach for space/spatial derivatives and advanced difference approach for time derivatives [15]. The following figure is an illustration of the Lax-Friedrichs scheme:

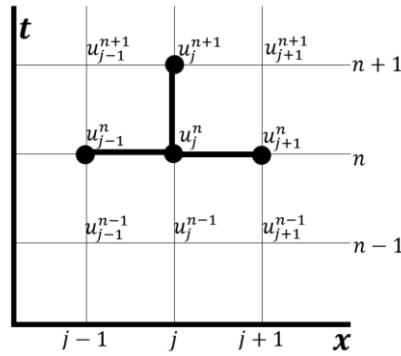


Figure 1. Illustration of the Lax-Friedrichs method

The solution begins with discretization of the differential equations in the dimensions of time and space/spatial. Discretization of the time and spatial variables of this three-dimensional equation aims to discretize the tsunami wave propagation equation (Equation 1-3) with boundary conditions $\frac{\partial \eta}{\partial x}(0,0,t) = \frac{\partial \eta}{\partial x}(a,b,t) = 0$ and $\frac{\partial \eta}{\partial y}(0,0,t) = \frac{\partial \eta}{\partial y}(a,b,t) = 0$. Discretization of Spatial Variables This three-dimensional equation uses a forward difference scheme, where the terms $\frac{\partial}{\partial x}\left(\frac{M^2}{D}\right)$, $\frac{\partial}{\partial y}\left(\frac{MN}{D}\right)$, $\frac{\partial}{\partial x}\left(\frac{MN}{D}\right)$, dan $\frac{\partial}{\partial y}\left(\frac{N^2}{D}\right)$ changed first using rules $\frac{\partial}{\partial x}\left(\frac{u}{v}\right) = \frac{\frac{\partial u}{\partial x}v - u\frac{\partial v}{\partial x}}{v^2}$ and $\frac{\partial}{\partial x}(wz) = \frac{\partial w}{\partial x}z + w\frac{\partial z}{\partial x}$. The results of the discretization of the time and spatial variables can be seen in equations 4-6 as follows:

$$\eta_{i,j}^{n+1} = \frac{1}{2}(\eta_{i+1,j+1}^n + \eta_{i-1,j-1}^n) - \frac{\Delta t}{2} \left(\frac{M_{i+1,j}^n - M_{i-1,j}^n}{\Delta x} + \frac{N_{i,j+1}^n - N_{i,j-1}^n}{\Delta y} \right) \quad (4)$$

$$M_{i,j}^{n+1} = \frac{1}{2}(M_{i+1,j+1}^n + M_{i-1,j-1}^n) \Delta t \left(\frac{M_{i,j}^n}{D_{i,j}^n} \left(2 \left(\frac{M_{i+1,j}^n - M_{i-1,j}^n}{2\Delta x} \right) - \frac{M_{i,j}^n}{D_{i,j}^n} \left(\frac{D_{i+1,j}^n - D_{i-1,j}^n}{2\Delta x} \right) + \left(\frac{N_{i,j+1}^n - N_{i,j-1}^n}{2\Delta x} \right) - \frac{N_{i,j}^n}{D_{i,j}^n} \left(\frac{D_{i,j+1}^n - D_{i,j-1}^n}{2\Delta y} \right) + \left(\frac{g n^2}{(D_{i,j}^n)^{\frac{4}{3}}} \sqrt{(M_{i,j}^n)^2 + (N_{i,j}^n)^2} \right) \right) + \frac{N_{i,j}^n}{2\Delta y D_{i,j}^n} (M_{i,j+1}^n - M_{i,j-1}^n) + \frac{g D_{i,j}^n}{2\Delta x} (\eta_{i+1,j}^n - \eta_{i-1,j}^n) \right) \quad (5)$$

$$N_{i,j}^{n+1} = \frac{1}{2}(N_{i+1,j+1}^n + N_{i-1,j-1}^n) - \Delta t \left(\frac{N_{i,j}^n}{D_{i,j}^n} \left(2 \left(\frac{N_{i,j+1}^n - N_{i,j-1}^n}{2\Delta y} \right) - \frac{M_{i,j}^n}{D_{i,j}^n} \left(\frac{D_{i+1,j}^n - D_{i-1,j}^n}{2\Delta x} \right) + \left(\frac{M_{i+1,j}^n - M_{i-1,j}^n}{2\Delta x} \right) - \frac{N_{i,j}^n}{D_{i,j}^n} \left(\frac{D_{i,j+1}^n - D_{i,j-1}^n}{2\Delta x} \right) + \left(\frac{g n^2}{(D_{i,j}^n)^{\frac{4}{3}}} \sqrt{(M_{i,j}^n)^2 + (N_{i,j}^n)^2} \right) \right) + \frac{M_{i,j}^n}{2\Delta x D_{i,j}^n} (N_{i+1,j}^n - N_{i-1,j}^n) + \frac{g D_{i,j}^n}{2\Delta x} (\eta_{i+1,j}^n - \eta_{i-1,j}^n) \right) \quad (6)$$

In the simulation using the Matlab program, the value of n is chosen from the average value of the roughness of the seabed materials. Simulation of the tsunami wave propagation model was carried out with an amplitude of 8 meters and a wave length of 50 km. The results of the numerical simulations show that the calculation results from the Lax-Friedrichs method obtain a fairly small difference compared to using manual calculations and the results of numerical simulations using the Runge Kutta Goemetric Mean (RKGM) method. This means that the Lax-Friedrichs method is a fairly good method for predicting the behavior of tsunami waves. The numerical simulation results of tsunami wave propagation using the Lax-Friedrichs method can be seen in Figure 2.

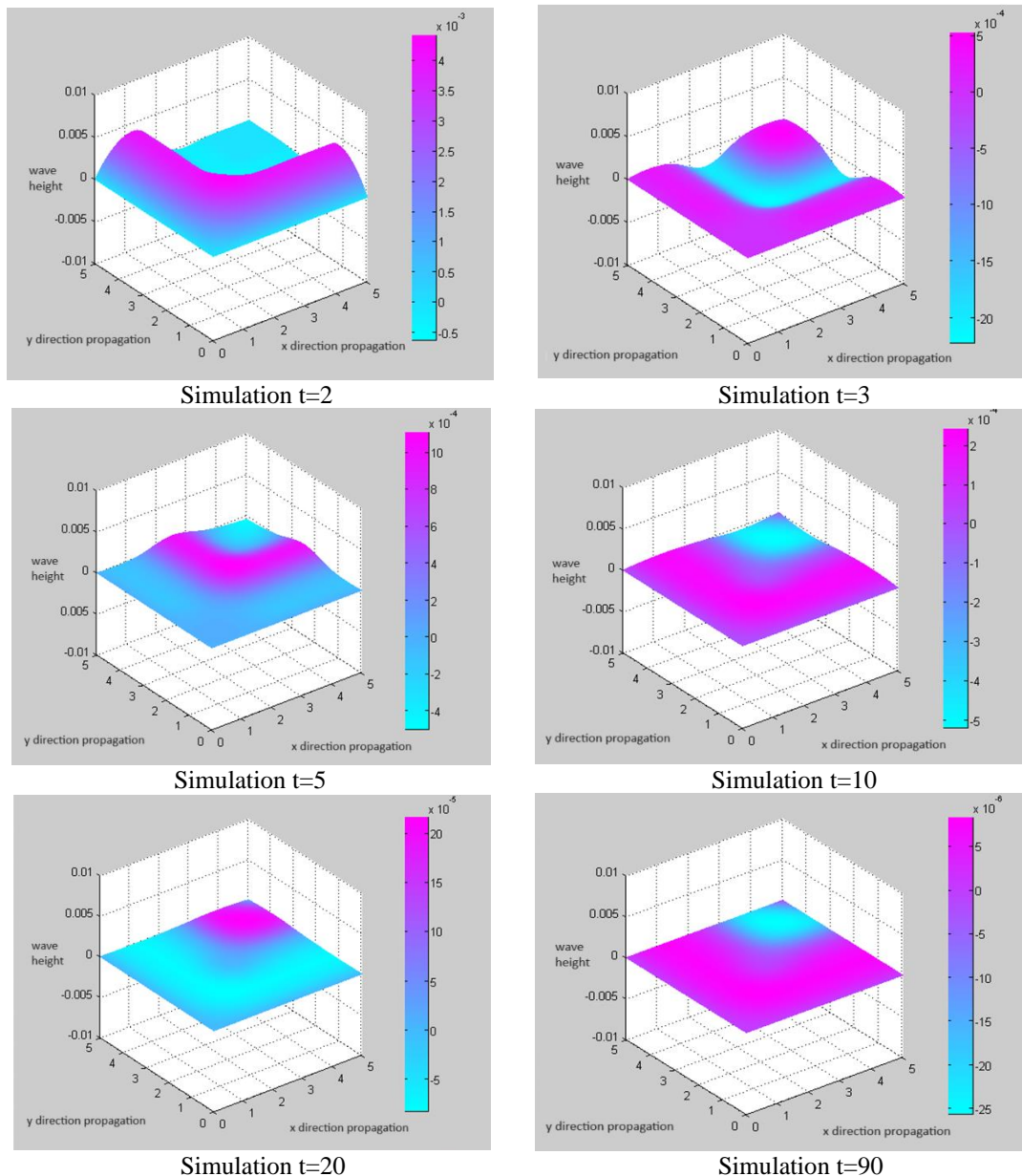


Figure 2. Simulation Result with Matlab

Numerical simulation using the Matlab program, obtained the results of the tsunami wave propagation plot as shown in Figure 2. Simulation of the tsunami wave at $t = 2$, the highest wave crest was about 4 meters and the valley formed was more than 0.5 meters. The height of the tsunami waves continued to decrease until $t = 10$, but there was an increase in the depth of the wave trough reaching more than 0.8 meters. At $t = 10$ to $t = 90$ the height of the tsunami waves fluctuates, when the waves rise, the depth of the valley decreases, when the valley gets deeper, the height of the tsunami waves decreases.

CONCLUSION

The Lax-Friedrichs method can be used to simulate tsunami wave propagation through the discretization process of the differential equations for tsunami wave propagation through the Finite Difference method approach. In general, tsunami waves move downward, but at certain times, the tsunami wave's rise. The most obvious changes in the height of the tsunami waves are seen only a few moments at the start of the

formation process. The height of the tsunami waves is directly proportional to the length of time it takes for the waves to calm down again.

REFERENCES

- [1] H. S. Naryanto, "Analisis Bahaya, Kerentanan Dan Risiko Bencana Tsunami Di Provinsi Papua Barat," *J. Alami*, vol. 3, no. 1, p. 10, 2019, doi: 10.29122/alami.v3i1.3399.
- [2] C. E. Synolakis, "The runup of long waves," California Institute of Technology, 1986.
- [3] E. T. Flouri, N. Kalligeris, G. Alexandrakakis, N. A. Kampanis, and C. E. Synolakis, "Application of a finite difference computational model to the simulation of earthquake generated tsunamis," *Appl. Numer. Math.*, vol. 67, pp. 111–125, 2013, doi: 10.1016/j.apnum.2011.06.003.
- [4] A. T. Layton and M. Van de Panne, "A numerically efficient and stable algorithm for animating water waves," *Vis. Comput.*, vol. 18, no. 1, pp. 41–53, 2002, doi: 10.1007/s003710100131.
- [5] F. Imamura, A. C. Yalçiner, and G. Ozyurt, *Tsunami modelling manual*. Tohoku, Jepang: Tohoku University, 2006.
- [6] W. Kongko and T. Schlurmann, "the Java Tsunami Model: Using Highly-Resolved Data To Model the Past Event and To Estimate the Future Hazard," *Proc. 32nd Conf. Coast. Eng.*, pp. 3670–3685, 2010, doi: 10.9753/icce.v32.management.25.
- [7] H. Latief, *Pedoman Teknik Pembuatan Peta Bahaya Rendaman Tsunami*. Bandung: Pusat Penelitian Mitigasi Bencana ITB, 2013.
- [8] S. Reese, B. A. Bradley, J. Bind, G. Smart, W. Power, and J. Sturman, "Empirical building fragilities from observed damage in the 2009 South Pacific tsunami," *Earth-Science Rev.*, vol. 107, no. 1–2, pp. 156–173, 2011, doi: 10.1016/j.earscirev.2011.01.009.
- [9] T. Riedlinger, H. Mehl, S. Dech, and J. Birkmann, "Tsunami risk assessment in Indonesia," *Nat. Hazards Earth Syst. Sci.*, vol. 11, pp. 67–82, 2011, doi: 10.5194/nhess-11-67-2011.
- [10] F. Usman, K. Murakami, and E. B. Kurniawan, "Study on Reducing Tsunami Inundation Energy by the Modification of Topography based on Local Wisdom," *Procedia Environ. Sci.*, vol. 20, pp. 642–650, 2014, doi: 10.1016/j.proenv.2014.03.077.
- [11] Y. Liu, Y. Shi, D. A. Yuen, E. O. D. Sevre, X. Yuan, and H. L. Xing, "Comparison of linear and nonlinear shallow wave water equations applied to tsunami waves over the China Sea," *Acta Geotech.*, vol. 4, no. 2, pp. 129–137, 2009, doi: 10.1007/s11440-008-0073-0.
- [12] M. Jamhuri, "Simulasi Perambatan Tsunami menggunakan Persamaan Gelombang Air-Dangkal," Malang, 2014.
- [13] Y. Fauzi, J. Rizal, F. Faisal, P. Novianti, and M. Sartika, "Pemodelan Penjalaran Gelombang Tsunami Melalui Pendekatan Finite Difference Method," in *Prosiding Seminar Nasional Matematika dan Pendidikan Matematika*, 2016, no. 1, pp. 17–22.
- [14] Y. Fauzi and Z. Memi Mayasari, "Theoretical study of the differential equation concept of tsunami waves using the leapfrog scheme," *J. Phys. Conf. Ser.*, vol. 1940, no. 1, 2021, doi: 10.1088/1742-6596/1940/1/012007.
- [15] L. Rezzolla, "Numerical methods for the solution of partial differential equations of fractional order," Albert Einstein Institute, 2003. doi: 10.1016/j.jcp.2003.07.008.

Optimal Control of Dengue Fever with the Influence of Vaccination and Insecticide Administration

Mariatul Qibtiah^{*(1)} and Pardi Affandi⁽¹⁾

⁽¹⁾Department of Mathematics, Faculty of Mathematics and Sciences, Lambung Mangkurat University, Loktabat Selatan, Banjarbaru, 70714, Indonesian

*Corresponding author: 2011011220004@mhs.ulm.ac.id

Abstract. Dengue Fever is a disease transmitted by vectors from mosquito bites that infect the Aedes Aegypti mosquito virus, this virus is transmitted from humans to vectors and from vectors to human hosts through female Aedes Aegypti mosquitoes. There are two living populations that play a role in the spread of DHF, namely humans, who are called hosts, and female Aedes Aegypti mosquitoes, which are known as carriers of the Aedes Aegypti mosquito virus. The purpose of this study is to explain the model of the spread of Dengue Fever, determine the disease-free and endemic equilibrium points, analyze the stability of the equilibrium points, and determine the optimal form of control. The method that will be used in this research is the Next Generation Matrix to determine the Basic Reproduction Number. Then, the optimal control solution is carried out using the Pontryagin Minimum Principle method. The result of this study is the establishment of the SEIR-SEI model for Dengue Fever. Based on this model, two equilibrium points are obtained, namely disease-free equilibrium points and endemic equilibrium points, and the basic matrix is obtained using the Next Generation Matrix method. Furthermore, analyzing the model, will make it asymptotically stable if $R_0 < 1$ at the disease-free equilibrium point and locally asymptotically stable if $R_0 > 1$ at the endemic equilibrium point. As well as forms of DHF control with policies provided including attack and administration of insecticides. Studies on the Dengue Fever disease model have been carried out with various types of analysis. One of them is research conducted by Puntani Pongsumpun, I-Ming Tang, and Napasool Wongvanich.

Keywords: Dengue fever; SEIR-SEI Models; Next Generation Matrix; Vaccination, and Insecticide Administration

INTRODUCTION

Indonesia is a country with a tropical climate and high rainfall. Therefore, the climate can cause various mosquito-borne diseases, one of which is Dengue Hemorrhagic Fever (DHF) caused by the Aedes Aegypti mosquito virus which is transmitted from human to human. Dengue virus transmission in mosquitoes occurs in two ways: horizontally and vertically. Dengue fever and dengue hemorrhagic fever can be caused by four serotypes of the virus: DEN-1, DEN-2, DEN-3, and DEN-4. Mosquitoes can become infected when they bite humans through horizontal transmission, but there is also vertical transmission through sexual contact with infected mosquitoes [1]. High fever, muscle and joint aches, pain behind the eyes, and a body rash that may go away and reappear are simple symptoms of dengue fever [1]. It is characterized by sudden onset of fever, which usually lasts for 2 to 7 days, and various nonspecific signs and symptoms. During the acute phase of the illness, it is difficult to distinguish DHF from dengue fever and other illnesses found in tropical areas [2]. Some individuals who have dengue fever can experience a more severe form of the disease. Severe bleeding can cause severe shock and death [1]. More than a hundred countries have been affected by this disease, including Indonesia [3].

Based on data from the Ministry of Health, in 2022, the number of DHF cases will reach 131,265 cases of which around 40% are children aged 0-14 years. Meanwhile, the number of deaths reached 1,135 cases with 73% occurring in children aged 0-14 years [4]. This figure has increased compared to 73,518 cases in 2021 with 705 deaths [5]. According to the work of Sabran & Jannah, dengue fever is widely developed in tropical and sub-tropical regions such as Indonesia. From 1968 to 2009, the World Health Organization (WHO) recorded Indonesia as the country with the highest number of dengue cases in Southeast Asia. The number of sufferers and the area of distribution is increasing along with the increase in population density [6].

According to research [1], many problems are related to vaccination for people infected with the Aedes Aegypti virus. One of them is the Dengvaxia vaccine. However, problems started to appear over time. On December 12, 2007, several countries stopped the program due to the need to control the spread of dengue fever. Two policies have been proposed to reduce cases of dengue virus infection. The first includes restoration

and application of insecticides, and the second involves isolation or quarantine and application of insecticides [1]. The main prevention that needs to be done to control the spread of dengue fever is to eradicate mosquitoes, which are vectors for spreading the disease. The presence of the Aedes Aegypti mosquito vector has a significant impact on the spread and number of cases of this disease. One way to do this is by spraying insecticides or cleaning the area of the house or where mosquitoes breed, especially the Aedes Aegypti mosquito. Thus, the infected human population can be reduced and the Aedes Aegypti mosquito population can also be reduced[3].

In addition, since no clinical studies have been conducted to determine whether quarantine of individuals infected with the dengue virus helps stop the epidemic, computer simulations facilitated by theoretical modeling and mathematical controls were developed by the researchers [7]. In the literature, several mathematical models have been proposed to explain how dengue virus spreads. To better understand how the disease is spread and controlled, a dengue epidemic model is needed. To describe the transmission of DHF, a mathematical model is often proposed. Esteva and Vargas (1999) suggested a model of dengue transmission in which the human population is constant but the mosquito populations vary. They provide a global analysis to ensure the global stability of the endemic equilibrium. Naowarat et al. (2011) suggested a dynamic model to measure the risk of dengue fever in humans. Pongsumpun and colleagues developed a mathematical model for dengue that incorporates additional aspects of the disease. Effects of the external incubation period and dynamic transmission are included. Modified Susceptible-Infected-Recovered (SIR) model using standard dynamic analysis [8]. In another study, the effect of the incubation period on viruses was studied in the SEIR model (Pongsumpun, 2017). The vector population in this model is divided into susceptible, infected, and contagious classes.

In this study, we will discuss dengue virus transmission again using the SEIR (Susceptible, Exposed, Infected, and Recovered) model which is used to hide the dynamics of the disease. Therefore, based on the problems above, our research solves the optimal control problem [1] which will discuss the model for the spread of dengue fever by determining the equilibrium point and analyzing disease-free and endemic stability. Then, the optimal control settlement is carried out by analyzing the first policy.

MATERIALS AND METHODS

Materials

a. Differential Equations

Definition 1 [9]

A linear ordinary differential equation of order n , in the dependent variable y and the independent variable x , is an equation that is in, or can be expressed in, the form

$$a_0(x) \frac{d^n y}{dx^n} + a_1(x) \frac{d^{n-1} y}{dx^{n-1}} + \dots + a_{n-1}(x) \frac{dy}{dx} + a_n(x)y = b(x) \quad (1)$$

Where a_0 is not identically zero.

b. Nonlinear Differential Equations

According to Braun (1992), the form of a system of nonlinear differential equations is as follows:

$$\begin{aligned} \frac{dx_1}{dt} &= f_1(x_1, x_2, \dots, x_n) \\ \frac{dx_2}{dt} &= f_2(x_1, x_2, \dots, x_n) \\ &\vdots \\ \frac{dx_n}{dt} &= f_n(x_1, x_2, \dots, x_n) \end{aligned} \quad (2)$$

where $f_i(x_1, x_2, \dots, x_n) \neq 0$ for every $i = 1, 2, \dots, n$ is a nonlinear function that is continuous on the interval with . The form of equation (2) can also be written in the following form:

$$\frac{dx}{dt} = f(x) \quad (3)$$

with,

$$\frac{dx}{dt} = \begin{pmatrix} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \\ \vdots \\ \frac{dx_n}{dt} \end{pmatrix}; f(x) = \begin{pmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_n(x) \end{pmatrix}; x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \text{ dan } f, x \in \mathbb{R}^n$$

c. *Equilibrium Point*

The equilibrium point is the point that states the equilibrium of a mathematical model. There are two kinds of this equilibrium point, namely, a disease-free equilibrium point is a condition where there is no longer any disease that attacks the population or it can be said that there are no individuals who are infected with the disease. While the disease endemic equilibrium point is a condition where the disease is always present in the population, meaning that there are always individuals who are infected with the disease.

Definition 2.

The point $\bar{x} \in \mathbb{R}^n$ is called the equilibrium point (equilibrium point) of the system (3) if $f(\bar{x}) = 0$.

d. *Basic Reproduction Number*

According to (Giesecke, 2017), there are several conditions that can arise, namely

1. If $R_0 < 1$ then the system will form stability at disease-free equilibrium, in other words the disease will disappear.
2. If $R_0 > 1$ then the disease-free equilibrium is unstable, in other words the disease will continue to grow into an epidemic.
3. If $R_0 = 1$ then the disease will persist or the disease will become endemic.

e. *Eigenvalues and Eigenvectors*

Definition 3.

If A is a matrix of $n \times n$, then a nonzero vector x on \mathbb{R}^n is called an eigenvector of A . If Ax is a scalar multiple of x such that,

$$Ax = \lambda x \quad (4)$$

Theorem. 1

If matrix A is of size $n \times n$, then λ is the eigenvalue of matrix B if and only if it satisfies the equation

$$|A - \lambda I| = 0 \quad (5)$$

with I matrix identity matrix. Equation (5) is called the characteristic equation of matrix A .

f. *Stability Analysis*

Theorem 2.

If λ_i it is the eigenvalue of the Jacobian matrix $n \times n$ at the equilibrium point \hat{x} and $Re(\lambda_i)$ is the real part of λ_i then:

1. If for every $i = 1, 2, 3, \dots, n, Re(\lambda_i) < 0$ then \hat{x} it is stable asymptotically.
2. If there is $Re(\lambda_i) > 0$ for a i then \hat{x} it is unstable

g. *Optimal Control*

State equation :

$$\dot{x} = f(x(t), u(t)) \quad (6)$$

with $x(t)$ state vector that models a state.

performance index:

$$J = S(x(t_f), t_f) + \int_{t_0}^{t_f} L(x(t), u(t), t) dt \quad (7)$$

with the boundary condition (initial state) $x(t_0) = x_0$ and t_f is the end time.

Here is the Lagrange form of performance index

$$J(u) = \int_{t_0}^{t_f} L(x(t), u(t), t) dt \quad (8)$$

h. Pontryagin Maximum Principle

According to Burghes (1980), here are the steps to solve the problem of optimal control based on the Pontryagin maximum principle:

1. Forms the Hamiltonian function from the performance index and state equation:

$$H(x(t), u(t), \lambda(t), t) = L(x(t), u(t), t) + \lambda(t)f(x(t), u(t), t) \quad (9)$$

2. Determine the costate equation:

$$\dot{\lambda} = -\frac{\partial H}{\partial x_i}, \quad i = 1, 2, \dots, n \quad (10)$$

3. Determine the state equation:

$$\dot{x} = -\frac{\partial H}{\partial \lambda_i}, \quad i = 1, 2, \dots, n \quad (11)$$

4. Define controls u :

$$\frac{\partial H}{\partial \lambda_i} = 0 \quad (12)$$

5. Substituting the controller $u^*(t)$ into the equation state to obtain the optimal system of equations.

Methods

The steps to be carried out in research regarding optimal control of DHF disease models are as follows:

1. This research is a literature study, starting from observing the high spread of dengue disease in Indonesia with the search for supporting theories in forming mathematical models.
2. Establishment of Dengue Fever Spread Model
In this section, the formation of a mathematical model of the dengue fever spread system is carried out. To understand the spread of Dengue Fever, certain assumptions are made so that compartment models can be made with Susceptible, Exposed, Infective, and Recovered subpopulations.
3. Determining the equilibrium point.
At this stage, we will find the disease-free equilibrium point ($I = 0$) and the endemic equilibrium point ($I \neq 0$) by substitution of the model equation. Next, the basic reproduction number will be determined by the Next Generation Matrix Method.
4. Analyzing Disease-Free and Endemic Equilibrium Points
At this stage, the Jacobian Matrix will be used by linearizing the model equation so that the Characteristic Value is obtained.
5. Method of completion of optimal control of Dengue hemorrhagic fever
Namely, by determining the model with control, namely with u_1 control, is a vaccination effort and u_2 is an insecticide.
6. Specifies the destination function (performance index) in the form of Lagrange.
Finding the optimal solution with the use of the Pontryagin Maximum principle (Hamiltonian Equation).
Determine the equation of state and costate
7. Making conclusions
The final stage is to draw conclusions on the discussion that has been carried out by writing this research article.

RESULTS AND DISCUSSION

1. The Mathematical Model of Dengue Fever

This mathematical model focuses on the spread of dengue fever using a system of non-linear ordinary differential equations. To simplify the model and still show the actual situation, to control the characteristics of dengue fever, a mathematical model was developed for human and mosquito vectors. Here, the SEIR model for the human population is considered. Here, the human population as a whole is divided into four classes: \widetilde{S}_h , \widetilde{E}_h , \widetilde{I}_h and \widetilde{R}_h . The mosquito vector population is also divided into three classes: \widetilde{S}_v , \widetilde{E}_v , \widetilde{I}_v . Below is a diagram of the transmission between the corresponding compartments.

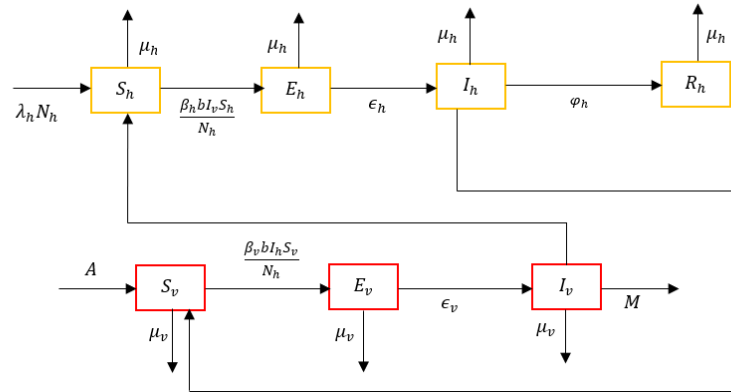


Figure 1. Dengue Disease Compartment

The variables used as the names of the compartments to construct the model can be seen in Table 1 below:

Table 1 Summary of description of variables and parameters for the measles model

Symbols	Description
\widetilde{S}_h	the number of susceptible human individuals at time t
\widetilde{E}_h	the number of exposed human individuals at time t
\widetilde{I}_h	the number of infected human individuals at time t
\widetilde{R}_h	the number of recovered human individuals at time t
\widetilde{S}_v	the number of susceptible vector population at time t
\widetilde{E}_v	the number of exposed vector population at time t
\widetilde{I}_v	the number of infected vector population at time t
A	the constant recruitment rate
M	the number of mosquitoes transovarially infected
N_h	the total number of human population
N_v	the total number of vector population

Model Assumptions

1. The population is uniform and mixes homogeneously. The total population size, $N_h(t) = \widetilde{S}_h(t) + \widetilde{E}_h(t) + \widetilde{I}_h(t) + \widetilde{R}_h(t)$ and $N_v(t) = \widetilde{S}_v(t) + \widetilde{E}_v(t) + \widetilde{I}_v(t)$ at any time $t > 0$

The transmission diagram duly admits the following system of differential equations defined:

$$\frac{d\widetilde{S}_h}{dt} = \lambda_h N_h - \frac{\beta_h b \widetilde{I}_v}{N_h} \widetilde{S}_h - \mu_h \widetilde{S}_h \quad (13)$$

$$\frac{d\widetilde{E}_h}{dt} = \frac{\beta_h b \widetilde{I}_v}{N_h} \widetilde{S}_h - \epsilon_h \widetilde{E}_h \quad (14)$$

$$\frac{d\widetilde{I}_h}{dt} = \epsilon_h \widetilde{E}_h - \mu_h \widetilde{I}_h - \varphi_h \widetilde{I}_h \quad (15)$$

$$\frac{d\widetilde{R}_h}{dt} = \varphi_h \widetilde{I}_h - \mu_h \widetilde{R}_h \quad (16)$$

$$\frac{d\widetilde{S}_v}{dt} = A - \frac{\beta_v b \widetilde{I}_h}{N_h} \widetilde{S}_v - \mu_v \widetilde{S}_v \quad (17)$$

$$\frac{d\widetilde{E}_v}{dt} = \frac{\beta_v b \widetilde{I}_h}{N_h} \widetilde{S}_v - \mu_v \widetilde{E}_v - \epsilon_v \widetilde{E}_v \quad (18)$$

$$\frac{d\widetilde{I}_v}{dt} = M + \epsilon_v \widetilde{E}_v - \mu_v \widetilde{I}_v \quad (19)$$

Assuming all parameter values are positive, it can be written as follows:

$$\lambda_h > 0, b > 0, \beta_h > 0, \mu_h > 0, \epsilon_h > 0, \varphi_h > 0, A > 0, \beta_v > 0, \epsilon_v > 0, M > 0$$

Similarly, the value of the function is $\mu_v(t)$ is $0 \leq \mu_v(t) \leq 1$.

To analyze the calculation, the model problem will be reviewed on the constant value of the human population $N_h(t)$ and the mosquito population $N_v(t)$. Can be written as:

$$N_h(t) = \widetilde{S}_h(t) + \widetilde{E}_h(t) + \widetilde{I}_h(t) + \widetilde{R}_h(t) \quad (20)$$

$$N_v(t) = \widetilde{S}_v(t) + \widetilde{E}_v(t) + \widetilde{I}_v(t) \quad (21)$$

Where:

λ_h	:	Per capita birth rate of the human population
μ_h	:	Death rate of the human population
b	:	Biting rate of the human population
μ_v	:	Death rate of the vector population
φ_h	:	Recovery rate of the vector population
β_h	:	Transmission probability of dengue virus from the vector population to the human population
β_v	:	Transmission probability of dengue virus from human population to vector
ϵ_h	:	Intrinsic incubation rate
ϵ_v	:	Extrinsic incubation rate

As a result of our model assuming that the total population of humans and vectors is constant, the rate of change of the human and vector populations becomes zero, i.e.:

$$\frac{d\widetilde{S}_h}{dt} + \frac{d\widetilde{E}_h}{dt} + \frac{d\widetilde{I}_h}{dt} + \frac{d\widetilde{R}_h}{dt} = 0 \quad (22)$$

$$\frac{d\widetilde{S}_v}{dt} + \frac{d\widetilde{E}_v}{dt} + \frac{d\widetilde{I}_v}{dt} = 0 \quad (23)$$

From the conditions above:

$$\lambda_h = \mu_h \quad (24)$$

$$N_v = \frac{A + M}{\mu_v} \quad (25)$$

Defining the normalized compartment variables from equations (13) – (21) as follows:

$$S_h = \frac{\widetilde{S}_h}{N_h}, \quad E_h = \frac{\widetilde{E}_h}{N_h}, \quad I_h = \frac{\widetilde{I}_h}{N_h}, \quad R_h = \frac{\widetilde{R}_h}{N_h} \quad (26)$$

$$S_v = \frac{\widetilde{S}_v}{N_v}, \quad E_v = \frac{\widetilde{E}_v}{N_v}, \quad I_v = \frac{\widetilde{I}_v}{N_v},$$

We also have:

$$S_h + E_h + I_h + R_h = 1 \quad (27)$$

$$S_v + E_v + I_v = 1 \quad (28)$$

The mathematical model of Equation (13) – (19) is now reduced to the following 5 Equations:

$$\frac{dS_h}{dt} = \mu_h(1 - S_h) - \frac{\beta_h b I_v}{N_h} S_h N_v \quad (29)$$

$$\frac{dE_h}{dt} = \frac{\beta_h b I_v}{N_h} S_h N_v - \mu_h E_h - \epsilon_h E_h \quad (30)$$

$$\frac{dI_h}{dt} = \epsilon_h E_h - \mu_h I_h - \varphi_h I_h \quad (31)$$

$$\frac{dE_v}{dt} = \beta_v b I_h S_v - \mu_v E_v - \epsilon_v E_v \quad (32)$$

$$\frac{dI_v}{dt} = \frac{M}{N_v} + \epsilon_v E_v - \mu_v I_v \quad (33)$$

2. Determining the Equilibrium Point of a Mathematical model

The Dengue hemorrhagic fever disease model has two equilibrium points, namely the disease-free equilibrium point and the endemic equilibrium point. The equilibrium point is determined by setting the right segment to zero.

$$\mu_h(1 - S_h) - \frac{\beta_h b I_v}{N_h} S_h N_v = 0$$

$$\frac{\beta_h b I_v}{N_h} S_h N_v - \mu_h E_h - \epsilon_h E_h = 0$$

$$\epsilon_h E_h - \mu_h I_h - \varphi_h I_h = 0$$

$$\beta_v b I_h S_v - \mu_v E_v - \epsilon_v E_v = 0$$

$$\frac{M}{N_v} + \epsilon_v E_v - \mu_v I_v = 0$$

When vertical transmission is possible, an equilibrium state can be obtained. According to [7] in his work found that the only possible equilibrium state is the endemic equilibrium state E_1 given by

$$E_1(S_h^{1*}, E_h^{1*}, I_h^{1*}, E_v^{1*}, I_v^{1*}) \quad (34)$$

where the equilibrium states $S_h^{1*}, E_h^{1*}, I_h^{1*}, E_v^{1*}, I_v^{1*}$ are given by Equations (24)–(28) in the work of Chanprasopchai et al.

For the system of absence of vertical transmission, the number of transovarially infected *M* mosquitoes is set to zero. With two equilibrium points, namely the disease-free equilibrium point and the endemic equilibrium point. In particular, the equilibrium point will occur at:

1) Disease-Free Equilibrium Point

A disease-free equilibrium point is a condition where no disease has struck or no individual is infected with the disease, using equations (13) – (19) so that:

$$\begin{aligned} E_h &= I_h = E_v = I_v = 0 \\ \mu_h(1 - S_h) - \frac{\beta_h b I_v}{N_h} S_h N_v &= 0 \\ \mu_h - \mu_h S_h - \frac{\beta_h b(0)}{N_h} S_h N_v &= 0 \\ \mu_h - \mu_h S_h &= 0 \\ S_h &= \frac{\mu_h}{\mu_h} = 1 \end{aligned}$$

So that

$$E_0 = (1, 0, 0, 0, 0) \quad (35)$$

2) Basic Reproduction Number

The basic reproduction number or commonly notated with R_0 is a quantity that expresses the level of transmission or spread of disease infection. To determine the basic reproduction number, Equation (13) – (19) is used with a disease-free equilibrium point of $E_0 = (S_h, E_h, I_h, E_v, I_v) = (1, 0, 0, 0, 0)$. The approach used to obtain using the Next Generation Matrix.

$$R_0^A = \frac{\beta_h \beta_v b^2 \epsilon_h \epsilon_v}{N_h \mu_v (\mu_h + \epsilon_h) (\mu_h + \varphi_h) (\mu_v + \epsilon_v)} \quad (36)$$

3) Endemic equilibrium point

$$E_2 = (S_h^{2*}, E_h^{2*}, I_h^{2*}, E_v^{2*}, I_v^{2*}) \quad (37)$$

Where

$$S_h^{2*} = \frac{N_h (\mu_v + \epsilon_v) (\beta_v b \epsilon_h \mu_h + \mu_v (\mu_h + \varphi_h) (\mu_h + \epsilon_h))}{b \beta_v \epsilon_h (b N_v \beta_h \epsilon_v + N_h \mu_h (\mu_v + \epsilon_v))} \quad (38)$$

$$E_h^{2*} = -\frac{\mu_h (-N_v \beta_h \beta_v b^2 \epsilon_h \epsilon_v + N_h \mu_v (\mu_h + \epsilon_h) (\mu_h + \varphi_h) (\mu_v + \epsilon_v))}{\beta_v b \epsilon_h (\mu_h + \epsilon_h) (b N_v \beta_h \epsilon_v + N_h \mu_h (\mu_v + \epsilon_v))} \quad (39)$$

$$I_h^{2*} = \frac{\mu_h (N_v \beta_h \beta_v b^2 \epsilon_h \epsilon_v - N_h \mu_v (\mu_h + \epsilon_h) (\mu_h + \varphi_h) (\mu_v + \epsilon_v))}{\beta_v b (\mu_h + \epsilon_h) (\mu_h + \varphi_h) (b N_v \beta_h \epsilon_v + N_h \mu_v (\mu_v + \epsilon_v))} \quad (40)$$

$$E_v^{2*} = -\frac{\mu_h \mu_v (-N_v \beta_h \beta_v b^2 \epsilon_h \epsilon_v + N_h \mu_v (\mu_h + \epsilon_h) (\mu_h + \varphi_h) (\mu_v + \epsilon_v))}{N_v \beta_h b \epsilon_v (\mu_v + \epsilon_v) (\beta_v b \epsilon_v \mu_h + \mu_v (\mu_h + \epsilon_h) (\mu_h + \varphi_h))} \quad (41)$$

$$I_v^{2*} = -\frac{\mu_h (-N_v \beta_h \beta_v b^2 \epsilon_h \epsilon_v + N_h \mu_v (\mu_h + \epsilon_h) (\mu_h + \varphi_h) (\mu_v + \epsilon_v))}{N_v \beta_v b (\mu_v + \epsilon_v) (b \beta_v \epsilon_v \mu_h + \mu_v (\mu_h + \epsilon_h) (\mu_h + \varphi_h))} \quad (42)$$

3. Local Stability Analysis

Stability at an equilibrium point can be determined based on the eigenvalue. Local stability in this model is divided into two, namely local stability at the disease-free equilibrium point and local stability at the endemic equilibrium point. The model of the spread of dengue disease in the system of Equations (29) – (33) is a system of nonlinear equations, so it is necessary to linearize the system using Jacobian matrix to analyze the stability of the equilibrium point.

$$J = \begin{bmatrix} -\left(\mu_h + \frac{\beta_h b I_v}{N_h} N_v\right) & 0 & 0 & 0 & -\frac{\beta_h b S_h}{N_h} N_v \\ \frac{\beta_h b I_v}{N_h} N_v & -(\mu_h + \epsilon_h) & 0 & 0 & \frac{\beta_h b S_h}{N_h} N_v \\ 0 & \delta_h & -(\mu_h + \varphi_h) & 0 & 0 \\ 0 & 0 & \beta_v b S_v & -(\mu_v + \epsilon_v) & 0 \\ 0 & 0 & 0 & \delta_v & -\mu_v \end{bmatrix}$$

so that a Jacobian matrix of dengue disease models is obtained which will be used to determine the stability of disease-free equilibrium points and the stability of endemic points.

a) stability of disease-free equilibrium points

The disease-free equilibrium E_0 is a local asymptotic stable equilibrium if: $R_0 < 1$.

Proof.

Substitution of values at the equilibrium point $E_0 = (1, 0, 0, 0, 0)$ so that the Jacobian matrix at the equilibrium point E_0 as follows:

$$J(E_0) = \begin{bmatrix} -\mu_h & 0 & 0 & 0 & -\frac{\beta_h b N_v}{N_h} \\ 0 & -k_1 & 0 & 0 & \frac{\beta_h b N_v}{N_h} \\ 0 & \delta_h & -k_2 & 0 & 0 \\ 0 & 0 & \beta_v b & -k_3 & 0 \\ 0 & 0 & 0 & \delta_v & -\mu_v \end{bmatrix}$$

Where: $k_1 = (\mu_h + \epsilon_h)$; $k_2 = (\mu_h + \varphi_h)$; $k_3 = (\mu_v + \epsilon_v)$

Thus, the Characteristic Equation obtained from the determinant equation:

$$\lambda^5 + e_1 \lambda^4 + e_2 \lambda^3 + e_3 \lambda^2 + e_4 \lambda + e_5 = 0$$

The above equation is contained in the work of [7] at proposition 2

The equation $\lambda^5 + e_1 \lambda^4 + e_2 \lambda^3 + e_3 \lambda^2 + e_4 \lambda + e_5 = 0$ is negative if RouthHurwitz's condition $\lambda_i > 0$ for $i = 1, 2, 3$

Because, $e_1, e_2, e_3, e_4, e_5 > 0$ and all parameters are positive, so that all conditions of Routh-Hurwitz are met, so the other eigenvalues are negative. Based on the results obtained, the disease-free equilibrium point is locally stable asymptotic with $R_0 < 1$.

b) stability of endemic equilibrium points

The equilibrium point state E_1 asymptotically stable when $R_0 > 1$

Proof.

Substitution of endemic equilibrium point values $E_2 = (S_h^{2*}, E_h^{2*}, I_h^{2*}, E_v^{2*}, I_v^{2*})$

$$J = \begin{bmatrix} -\left(\mu_h + \frac{\beta_h b I_v^{2*}}{N_h} N_v\right) & 0 & 0 & 0 & -\frac{\beta_h b S_h^{2*}}{N_h} N_v \\ \frac{\beta_h b I_v^{2*}}{N_h} N_v & -(\mu_h + \epsilon_h) & 0 & 0 & \frac{\beta_h b S_h^{2*}}{N_h} N_v \\ 0 & \delta_h & -(\mu_h + \varphi_h) & 0 & 0 \\ 0 & 0 & \beta_v b (1 - I_v^{2*} - E_v^{2*}) & -(\mu_v + \epsilon_v) & 0 \\ 0 & 0 & 0 & \delta_v & -\mu_v \end{bmatrix}$$

Local stability E_1 be formed through the linearization of equations (29) - (33) leading to determinant equations obtained earlier in the work [7]. To solve the eigenvalue equation, we get a similar characteristic equation, namely:

$$\lambda^5 + e_1 \lambda^4 + e_2 \lambda^3 + e_3 \lambda^2 + e_4 \lambda + e_5 = 0$$

The above equation is contained in the work of [7] at proposition 3

Because conditions $e_1, e_2, e_3, e_4, e_5 > 0$ and all positive parameters are met, the other eigenvalues are negative. Based on the above equation, the endemic equilibrium point E_2 is locally stable asymptotic with $R_0 > 1$.

4. Optimal Control

It is known that the DHF disease model is locally asymptotic stable at both obtained equilibrium points. Furthermore, because the model deployment is stable, the model will be given control. Control provided in the form of vaccination and insecticide administration. Both can be controlled by applying the principle of maximal pontryagin. The maximum principle of pontryagin is such a condition that optimal control completion can be obtained in accordance with the objective of maximizing the performance index. Optimal control of the model aims to maximize preventive efforts and the application of insecticides that can inhibit the transmission of dengue virus.

Model of the spread of dengue disease when given control (u_1, u_2) . The variable control $u_1(t)$ is vaccination in exposed and infected populations but vaccination has only a small effect so that it will not affect E_h and I_h . While variable $u_2(t)$ is control in mosquitoes where insecticides are given to eliminate or reduce mosquito populations and lead to assumptions later u_2 will work on E_v and I_v . Here is a model of dengue disease when given control.

$$\frac{d}{dt}S_h = \mu_h(1 - S_h) - \frac{\beta_h b I_v}{N_h} S_h N_v - u_1(t)S_h \quad (43)$$

$$\frac{d}{dt}E_h = \frac{\beta_h b I_v}{N_h} S_h N_v - \mu_h E_h - \delta_h E_h \quad (44)$$

$$\frac{d}{dt}I_h = \delta_h E_h - \mu_h I_h - \varphi_h I_h \quad (45)$$

$$\frac{d}{dt}E_v = \beta_v b I_h S_v - \mu_v E_v - \delta_v E_v - u_2(t)E_v \quad (46)$$

$$\frac{d}{dt}I_v = \frac{M}{N_v} + (\delta_v E_v - u_2(t))E_v - (\mu_v I_v + u_2(t))I_v \quad (47)$$

The performance index equation in the form of Lagrange for the DFH deployment model is as follows.

$$J(u_1, u_2) = \int_0^T \left(B_0 I_h(t) + \frac{1}{2} B_1 u_1^2 + \frac{1}{2} B_2 u_2^2 \right) dt \quad (48)$$

In the performance index equation for the DHF disease model with lagrange control, it is obtained as follows.

$$\begin{aligned} H &= L(x, u_1, u_2) + \lambda_1 \frac{dS_h}{dt} + \lambda_2 \frac{dE_h}{dt} + \lambda_3 \frac{dI_h}{dt} + \lambda_4 \frac{dE_v}{dt} + \lambda_5 \frac{dI_v}{dt} \quad (49) \\ H &= B_0 I_h(t) + \frac{1}{2} B_1 u_1^2 + \frac{1}{2} B_2 u_2^2 + \lambda_1 \left(\mu_h(1 - S_h) - \frac{b\beta_h}{N_h} S_h I_v N_v - u_1(t)S_h \right) + \\ &\quad \lambda_2 \left(\frac{b\beta_h}{N_h} S_h I_v N_v - (\delta_h + \mu_h)E_h \right) + \lambda_3 (\delta_h E_h - (\mu_h + r_h)I_h) + \lambda_4 (b\beta_v S_v I_h - \\ &\quad (\delta_v + \mu_v + u_2(t))E_v) + \lambda_5 \left(\frac{M}{N_v} + (\delta_v - u_2(t))E_v - (\mu_v + u_2(t))I_v \right) \end{aligned}$$

Next look for the costate equation of the model that has been given, obtained.

$$\lambda'_1 = -\frac{\partial H}{\partial S_h} = -\lambda_1 \left(-\mu_h - \frac{b\beta_h I_v N_v}{N_h} - u_1 \right) + \lambda_2 \frac{b\beta_h I_v N_v}{N_h} \quad (50)$$

$$\lambda'_2 = -\frac{\partial H}{\partial E_h} = -\lambda_2 (-\delta_h - \mu_h) - \lambda_3 \epsilon_h \quad (51)$$

$$\lambda'_3 = -\frac{\partial H}{\partial I_h} = -\lambda_3 (-\mu_h + r_h) - \lambda_4 b\beta_v S_v - B_0 \quad (52)$$

$$\lambda'_4 = -\frac{\partial H}{\partial E_v} = -\lambda_4 (-\epsilon_v + \mu_v + u_2) - \lambda_5 (\epsilon_v - u_2) \quad (53)$$

$$\lambda'_5 = -\frac{\partial H}{\partial I_v} = \left(\frac{b\beta_h S_h N_v}{N_h} \right) (\lambda_1 - \lambda_2) - \lambda_5 (-\mu_v - u_2) \quad (54)$$

Furthermore, obtaining the state equation is obtained as follows.

$$S'_h = \frac{\partial H}{\partial \lambda_1} = \mu_h(1 - S_h) - \frac{b\beta_h}{N_h} S_h I_v N_v - u_1(t)S_h \quad (55)$$

$$E'_h = \frac{\partial H}{\partial \lambda_2} = \frac{b\beta_h}{N_h} S_h I_v N_v - (\epsilon_h + \mu_h)E_h \quad (56)$$

$$I'_h = \frac{\partial H}{\partial \lambda_3} = \epsilon_h E_h - (\mu_h + \varphi_h)I_h \quad (57)$$

$$E'_v = \frac{\partial H}{\partial \lambda_4} = b\beta_v S_v I_h - (\epsilon_v + \mu_v + u_2(t))E_v \quad (58)$$

$$I'_v = \frac{\partial H}{\partial \lambda_5} = \frac{M}{N_v} + (\delta_v - u_2(t))E_v - (\mu_v + u_2(t))I_v \quad (59)$$

Then, look up the value of the u_1 control and u_2 . The control values u_1 and u_2 can be obtained as follows.

$$\frac{\partial H}{\partial u_1} = 0 \quad (60)$$

$$B_1 u_1 - S_h \lambda_1 = 0$$

$$u_1^* = \frac{S_h \lambda_1}{B_1}$$

$$\frac{\partial H}{\partial u_2} = 0 \quad (61)$$

$$B_2 u_2 - E_h \lambda_2 - E_v \lambda_4 - E_v \lambda_5 - I_v \lambda_5 = 0$$

$$B_2 u_2 = E_h \lambda_2 + E_v \lambda_4 + E_v \lambda_5 + I_v \lambda_5$$

$$u_2^* = \frac{E_h \lambda_2 + E_v \lambda_4 + E_v \lambda_5 + I_v \lambda_5}{B_2}$$

Thus, the optimal control variable is given by

$$u_1^*(t) = \max \left(\min \left(\frac{\lambda_1 S_h^*}{B_1}, u_1^{max} \right), 0 \right) \quad (62)$$

$$u_2^*(t) = \max \left(\min \left(\frac{E_h \lambda_2 + E_v \lambda_4 + E_v \lambda_5 + I_v \lambda_5}{B_2}, u_2^{max} \right), 0 \right) \quad (63)$$

Using the property of the control set, we can say that

$$u_1^* = \begin{cases} 0 & \text{if } \frac{\lambda_1 S_h}{B_1} \leq 0 \\ \frac{\lambda_1 S_h}{B_1} & \text{if } \frac{\lambda_1 S_h}{B_1} < u_1^{max} \\ u_1^{max} & \text{if } \frac{\lambda_1 S_h}{B_1} \geq u_1^{max} \end{cases} \quad (64)$$

$$u_2^* = \begin{cases} 0 & \text{if } \frac{E_h \lambda_2 + E_v \lambda_4 + E_v \lambda_5 + I_v \lambda_5}{B_2} \leq 0 \\ \frac{E_h \lambda_2 + E_v \lambda_4 + E_v \lambda_5 + I_v \lambda_5}{B_2} & \text{if } \frac{E_h \lambda_2 + E_v \lambda_4 + E_v \lambda_5 + I_v \lambda_5}{B_2} < u_2^{max} \\ u_2^{max} & \text{if } \frac{E_h \lambda_2 + E_v \lambda_4 + E_v \lambda_5 + I_v \lambda_5}{B_2} \geq u_2^{max} \end{cases} \quad (65)$$

Since the limit of u_1, u_2 is $0 \leq u_1, u_2 \leq 1$, it can be said that the limits of u_1^* and u_2^* are $0 \leq u_1^*, u_2^* \leq 1$. So that for control of vaccination obtained $u_1^* = \frac{S_h \lambda_1}{B_1}$ and control of insecticide spraying $u_2^* =$

$$\frac{E_h \lambda_2 + E_v \lambda_4 + E_v \lambda_5 + I_v \lambda_5}{B_2}$$

CONCLUSION

1. Based on the results and discussion it can be concluded
2. The mathematical model used in the spread of Dengue Hemorrhagic Fever is to use the SEI-EI model, as follows

$$\frac{dS_h}{dt} = \mu_h (1 - S_h) - \frac{\beta_h b I_v}{N_h} S_h N_v$$

$$\frac{dE_h}{dt} = \frac{\beta_h b I_v}{N_h} S_h N_v - \mu_h E_h - \epsilon_h E_h$$

$$\frac{dI_h}{dt} = \epsilon_h E_h - \mu_h I_h - \varphi_h I_h$$

$$\frac{dE_v}{dt} = \beta_v b I_h S_v - \mu_v E_v - \epsilon_v E_v$$

$$\frac{dI_v}{dt} = \frac{M}{N_v} + \epsilon_v E_v - \mu_v I_v$$

3. From the Dengue Fever disease model, a disease-free equilibrium point is obtained, namely,

$E_0 = (1, 0, 0, 0, 0)$ and the endemic equilibrium point

$E_2 = (S_h^{2*}, E_h^{2*}, I_h^{2*}, E_v^{2*}, I_v^{2*})$ in equations 38 - 42

4. The basic Reproduction Number used to obtain using the Next Generation Matrix.

$$R_0^4 = \frac{\beta_h \beta_v b^2 \epsilon_h \epsilon_v}{N_h \mu_v (\mu_h + \epsilon_h) (\mu_h + \varphi_h) (\mu_v + \epsilon_v)}$$

5. From the model obtained, it is found that the disease-free equilibrium point is asymptotically stable with the characteristic equation $\lambda^5 + e_1 \lambda^4 + e_2 \lambda^3 + e_3 \lambda^2 + e_4 \lambda + e_5 = 0$

6. The control values u_1 and u_2 are obtained as follows

$$u_1^* = \frac{S_h \lambda_1}{B_1}, u_2^* = \frac{E_h \lambda_2 + E_v \lambda_4 + E_v \lambda_5 + I_v \lambda_5}{B_2}$$

REFERENCES

- [1] Pongsumpun P, Tang I M and Wongvanich N 2019 Optimal Control of the Dengue Dynamical Transmission with Vertical Transmission *Adv. Differ. Equations* **2019**
- [2] Halstead S M 2017 Dengue and dengue hemorrhagic fever *Handb. Zoonoses, Second Ed. Sect. B Viral Zoonoses* **11** 89–99
- [3] Aziim S A Al and Arif D K 2022 Kontrol Optimal Penyebaran Penyakit Demam Berdarah dengan Pengaruh Penyemprotan Insektisida dan Pengobatan *J. Sains dan Seni ITS* **11**
- [4] Kementerian Kesehatan Republik Indonesia 2023 Data Kasus Demam Berdarah di Indonesia Tahun 2022 07 Februari 2023
- [5] Widi S (Data I id. 2022 Data Kasus DBD Tahun 2021 30 Agustus 2022 2
- [6] Sabran L O and Jannah M 2020 Model Matematika SEIRS-SEI pada Penyebaran Penyakit Demam Berdarah Dengue dengan Pengaruh Suhu *MAp (Mathematics Appl. J.* **2** 66–78
- [7] Chanprasopchai P, Tang I M and Pongsumpun P 2018 SIR Model for Dengue Disease with Effect of Dengue Vaccination *Comput. Math. Methods Med.* **2018**
- [8] Rahmalia D, Herlambang T, Rohmah A M and Muhith A 2020 Weights Optimization Using Firefly Algorithm on Optimal Control of Zika Disease from Dengue Symptoms by Vaccination *J. Phys. Conf. Ser.* **1594**
- [9] Shepley L R 2002 *Differential Equations and Their Solutions*

A Probabilistic Dynamic Programming Approach for Solving an Inventory Problem (Case study: Fertilizer Stock)

Ramya Rachmawati

Department of Mathematics, The Faculty of Mathematics and Natural Sciences,
The University of Bengkulu, Indonesia, e-mail ramya.rachmawati@unib.ac.id

*Corresponding author:

ORCHID IDs: <https://orcid.org/0009-0007-6094-9620>

Abstract. Probabilistic Dynamic Programming (PDP) is a powerful approach for solving complex optimization problems that involve uncertainties in making sequential decisions. This paper uses PDP to solve an inventory problem of fertilizers in a warehouse. The goal is to minimize the total cost of holding and ordering of inventory while meeting the demand for fertilizers. To apply PDP to inventory problems, a recursive equation that relates the optimal cost of managing the inventory over a planning horizon (from period $t + 1$ to T) needs to be formulated. Since the demand is uncertain and variable, the inventory level at the end of each period is a random variable that depends on the realization of the demand in that period. The optimal solution shows how much to order at each period, depending on the current inventory level and the future demand.

Keywords: Probabilistic Dynamic Programming, Inventory, Optimization, Sequential Decision-Making, Fertilizer Stock

INTRODUCTION

A company is often faced with inventory control problems which are relatively complicated. Too much inventory stored will result in risks arising from the existence of these inventories. These risks include the risk of loss of goods as well as deterioration or damage to goods when stored (Duan et al., 2018; Janssen et al., 2016).

In addition, stored goods often require extra care and high storage costs. On the other hand, too little inventory will result in the fulfillment of customer requests being delayed, which results in customers being disappointed or, even more fatal, the company will lose customers. Inventory problems are becoming increasingly complex, considering customer demand, in reality, is not known with certainty (Goltsos et al., 2022).

Inventory problems with a constant level of demand per unit of time can be solved using classical optimization, such as the Economic Order Quantity (EOQ) model (Busola et al., 2020; Oktavianty & Sukmono, 2020). However, this EOQ model cannot solve the inventory problem if the demand that occurs in each time period is not the same. For this reason, another approach is needed to solve the inventory problem with probabilistic and variable demand, one of which is the probabilistic dynamic programming approach.

Previous modeling efforts have been undertaken in inventory planning; for instance (Rizky, 2021) used dynamic programming to schedule the optimal production planning in an agro-industry company. (Kusumah & Ilmaniati, 2020) used dynamic programming to determine the optimal inventory level of subsidized and non-subsidized fertilizer. (Siddik & Harahap, 2019) used a string matching method to make a decision support tool for non-subsidized fertilizer inventory problems. (Oktavianty & Sukmono, 2020) used dynamic programming methods to determine the amount of raw materials in PT XYZ. They achieved more profits by using the dynamic programming method. (Widiyanesti & Christanto, 2018) utilized EOQ dan Production Order Quantity (POQ) to determine the optimal ordering policy of a rice-processing business in Kediri. (Marsetiya Utama et al., 2019) used a dynamic programming developed from the Wagner-Whitin algorithm to deal with multi-item inventory problems. They claim that the dynamic programming approach outperformed the EOQ models in their case study. (Wahyu Oktavia et al., 2021) applied EOQ method to determine the optimal quantity per order in PT UYZ. They claim that the model can reduce total inventory costs in the manufacturing company. (Veza et al., 2021) conducted research by utilizing the Monte Carlo method to determine the optimal fertilizer stock (case study in PT Tazar Guna Mandiri). In more recent research, (Fadhli et al., 2023) used dynamic programming to solve an inventory problem. The study aims to determine the quantity and cost of inventory of both organic and inorganic fertilizers that can meet farmer's demands.

METHODS

The research method in this paper is a literature study and case studies. The implementation model used in this paper is adapted from a textbook (Taha, 2013; Winston, 2022). The data used in large problems are randomized from real data taken from a fertilizer shop in Bengkulu, Indonesia. A probabilistic dynamic programming is utilized to solve the inventory problems. The computation experiment is done with Python programming language.

In probabilistic dynamic programming, the state in the following stages is not only determined by the current state and decisions but is also influenced by the distribution of opportunities for each subsequent stage. In general, the basic structure of probabilistic dynamic programming can be seen in Figure 1. At stage n , the process is in state s_n . For example, S is the set of states that may occur at stage $n + 1$. The system will move to state s_n with probability p_i if at stage n the system is in state s_n and decision x_n . C_i is the contribution from stage n to the objective function if the system goes to this state (Hillier & Lieberman, 2001).

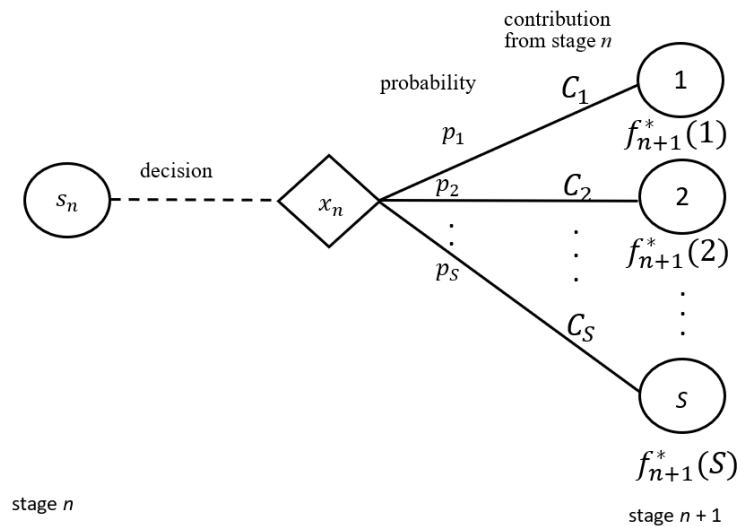


Figure 1. The Basic Structure of a Probabilistic Dynamic Programming

RESULTS AND DISCUSSION

In this section, first, a small test is presented, adapted from (Winston, 2022). Then the same method is applied to a larger problem.

A small test

A firm must determine how many units should be produced during the current period. Demand is a random variable. Each period's demand is equally likely to be 1 or 2 units. Production cost $c(x) = 3 + 2x$, $c(0)=0$. Production capacity = 4 units. The holding cost is \$1. The warehouse capacity is 3 units. Any inventory on hand at the end of period 3 can be sold at \$2 per unit. At the beginning of period 1, there is 1 unit of inventory. The firm wants to determine a production policy that minimizes the expected net cost incurred during the three periods.

Solution of the small test

First, define $f_t(i)$ to be the minimum expected net cost incurred during the periods $t, t + 1, \dots, 3$ when at the beginning of period t , the inventory level is i units. Then,

$$f_3(i) = \min_x \left\{ c(x) + \left(\frac{1}{2}\right)(i + x - 1) + \left(\frac{1}{2}\right)(i + x - 2) - \left(\frac{1}{2}\right)2(i + x - 1) - \left(\frac{1}{2}\right)2(i + x - 2) \right\}$$

where $x \in \{0, 1, 2, 3, 4\}$ and $(2 - i) \leq x \leq (4 - i)$.

The reason of that equation is as follows. If x units are produced during period 3, the net cost during that period is

$$(\text{expected production cost}) + (\text{expected holding cost}) - (\text{expected salvage value})$$

If x units are produced, the expected production cost is $c(x)$, and there is a $\frac{1}{2}$ chance that the period holding cost will be $i + x - 1$ and a $\frac{1}{2}$ chance that it will be $i + x - 2$. Meanwhile, the expected salvage value at the end of period 3 will be $\left(\frac{1}{2}\right) 2(i + x - 1) + \left(\frac{1}{2}\right) 2(i + x - 2)$. To make sure that period 3 demand is satisfied, we must have $i + x \geq 2$, and to make sure that period 3 inventory does not exceed warehouse capacity, we must have $i + x - 1 \leq 3$. Hence, we have $(2 - i) \leq x \leq (4 - i)$.

For $t = 1, 2$, the recursive relation is as follows

$$f_t(i) = \min_x \left\{ c(x) + \left(\frac{1}{2}\right)(i + x - 1) + \left(\frac{1}{2}\right)(i + x - 2) + \left(\frac{1}{2}\right)f_{t+1}(i + x - 1) + \left(\frac{1}{2}\right)f_{t+1}(i + x - 2) \right\}$$

where $x \in \{0, 1, 2, 3, 4\}$ and $(2 - i) \leq x \leq (4 - i)$.

The reasoning of that equation is as follows, for any month t production level x , the expected costs incurred during periods $t, t + 1, \dots, 3$ are the sums of the expected costs incurred during period t and the expected costs incurred during periods $t + 1, t + 2, \dots, 3$. The expected cost during month t will be $c(x) + \left(\frac{1}{2}\right)(i + x - 1) + \left(\frac{1}{2}\right)(i + x - 2)$. Here, no salvage value is received.

If x units are produced during month t , the expected cost during periods $t + 1, t + 2, \dots, 3$ is computed as follows. Half of the time, the demand during period t will be 1 unit, and the inventory at the beginning of period $t + 1$ will be $i + x - 1$. The expected costs incurred during periods $t + 1, t + 2, \dots, 3$ is $f_{t+1}(i + x - 1)$. Similarly, there is a $\frac{1}{2}$ chance that the inventory at the beginning of period $t + 1$ will be $i + x - 2$, so the expected costs incurred during periods $t + 1, t + 2, \dots, 3$ is $f_{t+1}(i + x - 2)$. In summary, the expected costs incurred during periods $t + 1, t + 2, \dots, 3$ is $\left(\frac{1}{2}\right)f_{t+1}(i + x - 1) + \left(\frac{1}{2}\right)f_{t+1}(i + x - 2)$.

$x_t(i)$ is period t production level attaining the minimum $f_t(i)$. We work backward until $f_1(1)$ is determined. The computations are summarized in Tables 1, 2 and 3.

Table 1. Stage 3 computations

i	x	$c(x)$	Expected holding cost	Expected salvage value	Total expected cost	$f_3(i)$ $x_3(i)$
3	0	0	3/2	3	-3/2*	$f_3(3) = -3/2$ $x_3(3) = 0$
3	1	5	5/2	5	5/2	
2	0	0	1/2	1	-1/2*	$f_3(2) = -1/2$ $x_3(2) = 0$
2	1	5	3/2	3	7/2	
2	2	7	5/2	5	9/2	
1	1	5	1/2	1	9/2*	$f_3(1) = 9/2$ $x_3(1) = 1$
1	2	7	3/2	3	11/2	
1	3	9	5/2	5	13/2	
0	2	7	1/2	1	13/2*	$f_3(0) = 13/2$ $x_3(0) = 2$
0	3	9	3/2	3	15/2	
0	4	11	5/2	5	17/2	

Table 2. Stage 2 computations

i	x	$c(x)$	Expected holding cost	Expected future cost	Total expected cost periods 2, 3	$f_2(i)$ $x_2(i)$
3	0	0	3/2	2	7/2*	$f_2(3) = 7/2$ $x_2(3) = 0$
3	1	5	5/2	-1	13/2	
2	0	0	1/2	11/2	6*	

i	x	$c(x)$	Expected holding cost	Expected future cost	Total expected cost periods 2, 3	$f_2(i)$ $x_2(i)$
2	1	5	3/2	2	17/2	$f_2(2) = 6$ $x_2(2) = 0$
2	2	7	5/2	-1	17/2	
1	1	5	1/2	11/2	11	$f_2(1) = 21/2$ $x_2(1) = 2$ or 3
1	2	7	3/2	2	21/2*	
1	3	9	5/2	-1	21/2*	
0	2	7	1/2	11/2	13	$f_2(0) = 25/2$ $x_2(0) = 3$ or 4
0	3	9	3/2	2	25/2*	
0	4	11	5/2	-1	25/2*	

Table 3. Stage 1 computations

i	x	$c(x)$	Expected holding cost	Expected future cost	Total expected cost periods 1 to 3	$f_1(i)$ $x_1(i)$
1	1	5	1/2	23/2	17	$f_1(1) = 65/4$ $x_1(1) = 3$
1	2	7	3/2	33/4	67/4	
1	3	9	5/2	19/4	65/4*	

We start by producing 3 units in period 1. Period 2's production level cannot be determined until period 1's demand is observed. Also, period 3's production level cannot be determined until period 2's demand is observed. As an illustration, suppose that period 1 and period 2 demands are both 1 unit. $x_1(1) = 3$ units will be produced during period 1. Then, period 2 will begin with $1 + 3 - 1 = 3$ units, and $x_2(3) = 0$ units will be produced during period 2. Then, period 3 will begin with $3 - 1 = 2$ units, and $x_3(2) = 0$ units will be produced during period 3.

A Case Study: Fertilizer Stock

Here, we apply the same methods as previous example, but in a larger problem. A fertilizer shop must determine how many units should be ordered during the current period. Demand is a random variable. Each period's demand with their probabilities can be seen in Table 4. Ordering cost $c(x) = 100000 + 15000x$, $c(0) = 0$. Ordering capacity = 4000 units. The holding cost is IDR 2700. The warehouse capacity is 3000 units. Any inventory on hand at the end of period 12 can be sold at IDR 30000 per unit. The firm would like to determine an ordering policy that minimizes the expected net cost incurred during the 12 periods.

Table 4. Demand of fertilizers (in sacks) and the probability of demand each month

Period	demand	probability	demand	probability
January	2000	0.5	1500	0.5
February	1500	0.5	1000	0.5
March	1500	0.5	2000	0.5
April	1500	0.6	1000	0.4
May	1500	0.6	1000	0.4
June	1500	0.7	1000	0.3
July	1500	0.8	1000	0.2
August	2000	0.5	1000	0.5
September	2000	0.6	1500	0.4
October	2000	0.5	1500	0.5
November	1500	0.7	1000	0.3
December	1500	0.6	1000	0.4

The computations are summarized in Tables 5, 6 and 7.

Table 5. Stage 12 computations of the fertilizers inventory problem

i	x	$c(x)$	expected holding cost	expected future cost	Total	$f_{12}(i)$ $x_{12}(i)$
0	2000	30100000	1890000.0	21000000.0	10990000.0	$f_{12}(0) = -13610000$
0	2500	37600000	3240000.0	36000000.0	4840000.0	$x_{12}(0) = 4000$
0	3000	45100000	4590000.0	51000000.0	-1310000.0	
0	3500	52600000	5940000.0	66000000.0	-7460000.0	
0	4000	60100000	7290000.0	81000000.0	-13610000.0	
500	1500	22600000	1890000.0	21000000.0	3490000.0	$f_{12}(500) = -21110000$
500	2000	30100000	3240000.0	36000000.0	-2660000.0	$x_{12}(500) = 3500$
500	2500	37600000	4590000.0	51000000.0	-8810000.0	
500	3000	45100000	5940000.0	66000000.0	-14960000.0	
500	3500	52600000	7290000.0	81000000.0	-21110000.0	
1000	1000	15100000	1890000.0	21000000.0	-4010000.0	$f_{12}(1000) = -$ 21110000
1000	1500	22600000	3240000.0	36000000.0	-10160000.0	
1000	2000	30100000	4590000.0	51000000.0	-16310000.0	$x_{12}(1000) = 3000$
1000	2500	37600000	5940000.0	66000000.0	-22460000.0	
1000	3000	45100000	7290000.0	81000000.0	-28610000.0	
1500	500	7600000	1890000.0	21000000.0	-11510000.0	$f_{12}(1500) = -$ 36110000
1500	1000	15100000	3240000.0	36000000.0	-17660000.0	
1500	1500	22600000	4590000.0	51000000.0	-23810000.0	$x_{12}(1500) = 2500$
1500	2000	30100000	5940000.0	66000000.0	-29960000.0	
1500	2500	37600000	7290000.0	81000000.0	-36110000.0	
2000	0	0	1890000.0	21000000.0	-19110000.0	$f_{12}(2000) = -$ 43610000
2000	500	7600000	3240000.0	36000000.0	-25160000.0	
2000	1000	15100000	4590000.0	51000000.0	-31310000.0	$x_{12}(2000) = 2000$
2000	1500	22600000	5940000.0	66000000.0	-37460000.0	
2000	2000	30100000	7290000.0	81000000.0	-43610000.0	
2500	0	0	3240000.0	36000000.0	-32760000.0	$f_{12}(2500) = -$ 51110000
2500	500	7600000	4590000.0	51000000.0	-38810000.0	
2500	1000	15100000	5940000.0	66000000.0	-44960000.0	$x_{12}(2500) = 1500$
2500	1500	22600000	7290000.0	81000000.0	-51110000.0	
3000	0	0	4590000.0	51000000.0	-46410000.0	$f_{12}(3000) = -$ 58610000
3000	500	7600000	5940000.0	66000000.0	-52460000.0	
3000	1000	15100000	7290000.0	81000000.0	-58610000.0	$x_{12}(3000) = 1000$
3500	0	0	5940000.0	66000000.0	-60060000.0	$f_{12}(3500) = -$ 66110000
3500	500	7600000	7290000.0	81000000.0	-66110000.0	$x_{12}(3500) = 500$

Table 6. Stage 11 computations of the fertilizers inventory problem

i	x	$c(x)$	expected holding cost	expected future cost	Total	$f_{11}(i)$ $x_{11}(i)$
0	2000	30100000	1755000.0	-23360000.0	8495000.0	$f_{11}(0) = 8495000.0$
0	2500	37600000	3105000.0	-30860000.0	9845000.0	$x_{11}(0) = 2000$
0	3000	45100000	4455000.0	-38360000.0	11195000.0	
0	3500	52600000	5805000.0	-45860000.0	12545000.0	
0	4000	60100000	7154999.9	-53360000.0	13895000.0	
500	1500	22600000	1755000.0	-23360000.0	995000.0	$f_{11}(500) = 995000.0$
500	2000	30100000	3105000.0	-30860000.0	2345000.0	$x_{11}(500) = 1500$
500	2500	37600000	4455000.0	-38360000.0	3695000.0	
500	3000	45100000	5805000.0	-45860000.0	5045000.0	
500	3500	52600000	7154999.9	-53360000.0	6395000.0	
1000	1000	15100000	1755000.0	-23360000.0	-6504999.9	$f_{11}(1000) = -6504999.9$
1000	1500	22600000	3105000.0	-30860000.0	-5155000.0	$x_{11}(1000) = 1000$
1000	2000	30100000	4455000.0	-38360000.0	-3805000.0	
1000	2500	37600000	5805000.0	-45860000.0	-2454999.9	
1000	3000	45100000	7154999.9	-53360000.0	-1105000.0	
1500	500	7600000	1755000.0	-23360000.0	-14004999.9	$f_{11}(1500) = -14004999.9$
1500	1000	15100000	3105000.0	-30860000.0	-12655000.0	$x_{11}(1500) = 500$
1500	1500	22600000	4455000.0	-38360000.0	-11305000.0	
1500	2000	30100000	5805000.0	-45860000.0	-9954999.9	
1500	2500	37600000	7154999.9	-53360000.0	-8605000.0	

i	x	$c(x)$	expected holding cost	expected future cost	Total	$f_i(i)$ $x_1(i)$
2000	0	0	1755000.0	-23360000.0	-21605000.0	$f_i(2000) = -21605000.0$
2000	500	7600000	3105000.0	-30860000.0	-20155000.0	$x_1(2000) = 0$
2000	1000	15100000	4455000.0	-38360000.0	-18805000.0	
2000	1500	22600000	5805000.0	-45860000.0	-17454999.9	
2000	2000	30100000	7154999.9	-53360000.0	-16105000.0	
2500	0	0	3105000.0	-30860000.0	-27755000.0	$f_i(2500) = -27755000.0$
2500	500	7600000	4455000.0	-38360000.0	-26305000.0	$x_1(2500) = 0$
2500	1000	15100000	5805000.0	-45860000.0	-24954999.9	
2500	1500	22600000	7154999.9	-53360000.0	-23605000.0	
3000	0	0	4455000.0	-38360000.0	-33905000.0	$f_i(3000) = -33905000.0$
3000	500	7600000	5805000.0	-45860000.0	-32454999.9	$x_1(3000) = 0$
3000	1000	15100000	7154999.9	-53360000.0	-31105000.0	
3500	0	0	5805000.0	-45860000.0	-40055000.0	$f_i(3500) = -40055000.0$
3500	500	7600000	7154999.9	-53360000.0	-38605000.0	$x_1(3500) = 0$

Table 7. Stage 1 computations of the fertilizers inventory problem

i	x	$c(x)$	expected holding cost	expected future cost	Total	$f_i(i)$ $x_1(i)$
0	2000	30100000	675000.0	222140000.0	252915000.0	$f_i(0) = 252915000$
0	2500	37600000	2025000.0	214640000.0	254265000.0	$x_1(0) = 2000$
0	3000	45100000	3375000.0	207140000.0	255615000.0	
0	3500	52600000	4725000.0	199590000.0	256915000.0	
0	4000	60100000	6075000.0	192715000.0	258890000.0	
500	1500	22600000	675000.0	222140000.0	245415000.0	$f_i(500) = 245415000$
500	2000	30100000	2025000.0	214640000.0	246765000.0	$x_1(500) = 1500$
500	2500	37600000	3375000.0	207140000.0	248115000.0	
500	3000	45100000	4725000.0	199590000.0	249415000.0	
500	3500	52600000	6075000.0	192715000.0	251390000.0	
1000	1000	15100000	675000.0	222140000.0	237915000.0	$f_i(1000) = 237915000$
1000	1500	22600000	2025000.0	214640000.0	239265000.0	$x_1(1000) = 1000$
1000	2000	30100000	3375000.0	207140000.0	240615000.0	
1000	2500	37600000	4725000.0	199590000.0	241915000.0	
1000	3000	45100000	6075000.0	192715000.0	243890000.0	
1500	500	7600000	675000.0	222140000.0	230415000.0	$f_i(1500) = 230415000$
1500	1000	15100000	2025000.0	214640000.0	231765000.0	$x_1(1500) = 500$
1500	1500	22600000	3375000.0	207140000.0	233115000.0	
1500	2000	30100000	4725000.0	199590000.0	234415000.0	
1500	2500	37600000	6075000.0	192715000.0	236390000.0	
2000	0	0	675000.0	222140000.0	222815000.0	$f_i(2000) = 222815000$
2000	500	7600000	2025000.0	214640000.0	224265000.0	$x_1(2000) = 0$
2000	1000	15100000	3375000.0	207140000.0	225615000.0	
2000	1500	22600000	4725000.0	199590000.0	226915000.0	
2000	2000	30100000	6075000.0	192715000.0	228890000.0	
2500	0	0	2025000.0	214640000.0	216665000.0	$f_i(2500) = 216665000$
2500	500	7600000	3375000.0	207140000.0	218115000.0	$x_1(2500) = 0$
2500	1000	15100000	4725000.0	199590000.0	219415000.0	
2500	1500	22600000	6075000.0	192715000.0	221390000.0	
3000	0	0	3375000.0	207140000.0	210515000.0	$f_i(3000) = 210515000$
3000	500	7600000	4725000.0	199590000.0	211915000.0	$x_1(3000) = 0$
3000	1000	15100000	6075000.0	192715000.0	213890000.0	
3500	0	0	4725000.0	199590000.0	204315000.0	$f_i(3500) = 204315000$
3500	500	7600000	6075000.0	192715000.0	206390000.0	$x_1(3500) = 0$

We intentionally only present the Stage 12, 11 and 1 computations for the sake of simplicity of the presentation of this article. However, we offer the optimal solutions for each stage as follows.

Table 8. The optimal solution of each stage of the fertilizers inventory problem

$f_1(i)$	$f_2(i)$	$f_3(i)$	$f_4(i)$	$f_5(i)$	$f_6(i)$	$f_7(i)$	$f_8(i)$	$f_9(i)$	$f_{10}(i)$	$f_{11}(i)$	$f_{12}(i)$
$x_1(i)$	$x_2(i)$	$x_3(i)$	$x_4(i)$	$x_5(i)$	$x_6(i)$	$x_7(i)$	$x_8(i)$	$x_9(i)$	$x_{10}(i)$	$x_{11}(i)$	$x_{12}(i)$
$f_1(0) = 252915000$	$f_2(0) = 225890000$	$f_3(0) = 205015000$	$f_4(0) = 177990000$	$f_5(0) = 156500000$	$f_6(0) = 135010000$	$f_7(0) = 112905000$	$f_8(0) = 90185000$	$f_9(0) = 63160000$	$f_{10}(0) = 35520000$	$f_{11}(0) = 8495000$	$f_{12}(0) = 13610000$
$x_1(0) = 2000$	$x_2(0) = 2000$	$x_3(0) = 2000$	$x_4(0) = 2000$	$x_5(0) = 2000$	$x_6(0) = 2000$	$x_7(0) = 2000$	$x_8(0) = 2000$	$x_9(0) = 2000$	$x_{10}(0) = 2000$	$x_{11}(0) = 2000$	$x_{12}(0) = 4000$
$f_1(500) = 245415000$	$f_2(500) = 218390000$	$f_3(500) = 197515000$	$f_4(500) = 170490000$	$f_5(500) = 149000000$	$f_6(500) = 127510000$	$f_7(500) = 105405000$	$f_8(500) = 82685000$	$f_9(500) = 55660000$	$f_{10}(500) = 28020000$	$f_{11}(500) = 995000$	$f_{12}(500) = 21110000$
$x_1(500) = 1500$	$x_2(500) = 1500$	$x_3(500) = 1500$	$x_4(500) = 1500$	$x_5(500) = 1500$	$x_6(500) = 1500$	$x_7(500) = 1500$	$x_8(500) = 1500$	$x_9(500) = 1500$	$x_{10}(500) = 1500$	$x_{11}(500) = 1500$	$x_{12}(500) = 3500$
$f_1(1000) = 237915000$	$f_2(1000) = 210890000$	$f_3(1000) = 190015000$	$f_4(1000) = 162990000$	$f_5(1000) = 141500000$	$f_6(1000) = 120010000$	$f_7(1000) = 97905000$	$f_8(1000) = 75185000$	$f_9(1000) = 48160000$	$f_{10}(1000) = 20520000$	$f_{11}(1000) = 6504999$	$f_{12}(1000) = 21110000$
$x_1(1000) = 1000$	$x_2(1000) = 1000$	$x_3(1000) = 1000$	$x_4(1000) = 1000$	$x_5(1000) = 1000$	$x_6(1000) = 1000$	$x_7(1000) = 1000$	$x_8(1000) = 1000$	$x_9(1000) = 1000$	$x_{10}(1000) = 1000$	$x_{11}(1000) = 1000$	$x_{12}(1000) = 3000$
$f_1(1500) = 230415000$	$f_2(1500) = 203390000$	$f_3(1500) = 182515000$	$f_4(1500) = 155490000$	$f_5(1500) = 134000000$	$f_6(1500) = 112510000$	$f_7(1500) = 90405000$	$f_8(1500) = 67685000$	$f_9(1500) = 40660000$	$f_{10}(1500) = 13020000$	$f_{11}(1500) = 14004999$	$f_{12}(1500) = 36110000$
$x_1(1500) = 500$	$x_2(1500) = 500$	$x_3(1500) = 500$	$x_4(1500) = 500$	$x_5(1500) = 500$	$x_6(1500) = 500$	$x_7(1500) = 500$	$x_8(1500) = 500$	$x_9(1500) = 500$	$x_{10}(1500) = 500$	$x_{11}(1500) = 500$	$x_{12}(1500) = 2500$
$f_1(2000) = 222815000$	$f_2(2000) = 195790000$	$f_3(2000) = 174915000$	$f_4(2000) = 147890000$	$f_5(2000) = 126400000$	$f_6(2000) = 104910000$	$f_7(2000) = 82805000$	$f_8(2000) = 60085000$	$f_9(2000) = 33060000$	$f_{10}(2000) = 5420000$	$f_{11}(2000) = 21605000$	$f_{12}(2000) = 43610000$
$x_1(2000) = 0$	$x_2(2000) = 0$	$x_3(2000) = 0$	$x_4(2000) = 0$	$x_5(2000) = 0$	$x_6(2000) = 0$	$x_7(2000) = 0$	$x_8(2000) = 0$	$x_9(2000) = 0$	$x_{10}(2000) = 0$	$x_{11}(2000) = 0$	$x_{12}(2000) = 2000$
$f_1(2500) = 216665000$	$f_2(2500) = 189640000$	$f_3(2500) = 168765000$	$f_4(2500) = 141740000$	$f_5(2500) = 120250000$	$f_6(2500) = 98760000$	$f_7(2500) = 76655000$	$f_8(2500) = 53935000$	$f_9(2500) = 26910000$	$f_{10}(2500) = 729999$	$f_{11}(2500) = 27755000$	$f_{12}(2500) = 51110000$
$x_1(2500) = 0$	$x_2(2500) = 0$	$x_3(2500) = 0$	$x_4(2500) = 0$	$x_5(2500) = 0$	$x_6(2500) = 0$	$x_7(2500) = 0$	$x_8(2500) = 0$	$x_9(2500) = 0$	$x_{10}(2500) = 0$	$x_{11}(2500) = 0$	$x_{12}(2500) = 1500$
$f_1(3000) = 210515000$	$f_2(3000) = 183440000$	$f_3(3000) = 162615000$	$f_4(3000) = 135550000$	$f_5(3000) = 114060000$	$f_6(3000) = 92580000$	$f_7(3000) = 70485000$	$f_8(3000) = 47785000$	$f_9(3000) = 20760000$	$f_{10}(3000) = 6879999$	$f_{11}(3000) = 33905000$	$f_{12}(3000) = 58610000$
$x_1(3000) = 0$	$x_2(3000) = 0$	$x_3(3000) = 0$	$x_4(3000) = 0$	$x_5(3000) = 0$	$x_6(3000) = 0$	$x_7(3000) = 0$	$x_8(3000) = 0$	$x_9(3000) = 0$	$x_{10}(3000) = 0$	$x_{11}(3000) = 0$	$x_{12}(3000) = 1000$
$f_1(3500) = 204315000$	$f_2(3500) = 177915000$	$f_3(3500) = 156415000$	$f_4(3500) = 129880000$	$f_5(3500) = 108390000$	$f_6(3500) = 86765000$	$f_7(3500) = 64525000$	$f_8(3500) = 41585000$	$f_9(3500) = 14570000$	$f_{10}(3500) = 13080000$	$f_{11}(3500) = 40055000$	$f_{12}(3500) = 66110000$
$x_1(3500) = 0$	$x_2(3500) = 0$	$x_3(3500) = 0$	$x_4(3500) = 0$	$x_5(3500) = 0$	$x_6(3500) = 0$	$x_7(3500) = 0$	$x_8(3500) = 0$	$x_9(3500) = 0$	$x_{10}(3500) = 0$	$x_{11}(3500) = 0$	$x_{12}(3500) = 500$

CONCLUSION

Based on the model demonstration, it can be concluded that the inventory system problem with uncertain demand can be solved through a probabilistic dynamic programming approach. This paper is limited to relatively short planning periods, namely three and 12 periods. For future research, case studies need to be conducted for a more extended planning period and involve more than one type of item.

REFERENCES

- [1]. Busola, K., Olaleke, O., Adegbuyi, O., Omotayo, A., & Ibidunni, S. (2020). Article ID: IJM_11_07_074 Economic Performance using ABC and EOQ Models. *International Journal of Management (IJM)*, 11(7), 835–848. <https://doi.org/10.34218/IJM.11.7.2020.074>
- [2]. Duan, Y., Cao, Y., & Huo, J. (2018). Optimal pricing, production, and inventory for deteriorating items under demand uncertainty: The finite horizon case. *Applied Mathematical Modelling*, 58, 331–348. <https://doi.org/10.1016/j.apm.2018.02.004>
- [3]. Fadhli, R., Isnaini Hadiyul Umam, M., Nur, M., & Islam Negeri Sultan Syarif Kasim Riau, U. (2023). Analisis Persediaan Pupuk Anorganik dan Organik Menggunakan Metode Pemrograman Dinamis. *Journal of Information Technology and Computer Science (INTECOMS)*, 6(1).
- [4]. Goltos, T. E., Syntetos, A. A., Glock, C. H., & Ioannou, G. (2022). Inventory – forecasting: Mind the gap. In *European Journal of Operational Research* (Vol. 299, Issue 2, pp. 397–419). Elsevier B.V. <https://doi.org/10.1016/j.ejor.2021.07.040>
- [5]. Hillier, F. S., & Lieberman, G. J. (2001). *Introduction to operations research*. McGrawHill.
- [6]. Janssen, L., Claus, T., & Sauer, J. (2016). Literature review of deteriorating inventory models by key topics from 2012 to 2015. In *International Journal of Production Economics* (Vol. 182, pp. 86–112). Elsevier B.V. <https://doi.org/10.1016/j.ijpe.2016.08.019>
- [7]. Kusumah, R. T., & Ilmaniati, A. (2020). Optimasi Persediaan Pupuk Non-Subsidi Menggunakan Program Dinamis Model Inventory (UD. Barokah). *Jurnal Media Teknik Dan Sistem Industri*, 3(2), 67. <https://doi.org/10.35194/jmts.v3i2.767>
- [8]. Marsetiya Utama, D., Garside, A. K., & Pamungkas, H. (2019). Model Program Dinamis untuk Lot Size Multi Item dengan Kendala Kapasitas Gudang. *Jurnal Teknik Industri*, 14(1).
- [9]. Oktaviany, V. N., & Sukmono, T. (2020). Optimalisasi Penentuan Biaya Minimum pada Pengendalian Persediaan Bahan Baku dengan Menggunakan Metode Dynamic Programming (Studi Kasus di PT XYZ). *Spektrum Industri*, 18(1), 15. <https://doi.org/10.12928/si.v18i1.10972>
- [10]. Rizky, A. N. (2021). Program Dinamik Pada Perencanaan Produksi Dan Pengendalian Persediaan PT Ganesha Abaditama. *Jurnal Optimasi Teknik Industri*, 03(01), 14–18.
- [11]. Siddik, M., & Harahap, S. Z. (2019). Aplikasi Pendukung Keputusan Pupuk Non Subsidi Dengan Metode String Matching (Studi Kasus Cv. Family Groups Labuhanbatu Selatan). *U-NET Jurnal Teknik Informatika*, 3(1), 14–20. <https://doi.org/10.52332/u-net.v3i1.17>
- [12]. Taha, H. A. (2013). *Operations research: an introduction*. Pearson Education India.
- [13]. Veza, O., Illya Badri, R., & kunci-Monte Carlo, K. (2021). SIMULASI PENGENDALIAN PERSEDIAAN PUPUK MENGGUNAKAN METODE MONTE CARLO. *Engineering And Technology International Journal Juli*, 3(2), 2714–2755. <https://doi.org/10.55642/eatij.v3i02>
- [14]. Wahyu Oktavia, C., Natalia, C., Jaya, A., Raya Benowo No, J., & Raya Cisauk, J. (2021). Analisis Pengaruh Pendekatan Economic Order Quantity terhadap Penghematan Biaya Persediaan. *XV(1)*, 103–117.
- [15]. Widiyanesti, S., & Christanto, M. D. (2018). Analisis Optimasi Biaya Bahan Baku melalui Implementasi Metode Lot Sizing pada Beras Cap Koi UD. Jongbiru Kediri. *E-Proceeding of Management*, 5(3).
- [16]. Winston, W. L. (2022). *Operations research: applications and algorithms*. Cengage Learning.

Application of the Arima Method for Forecasting Coffee Production in South Kalimantan Province

Regina Dietya Erika⁽¹⁾, Nur Salam*⁽¹⁾, Aprida Siska Lestia⁽²⁾

⁽¹⁾Statistics Study Program, FMIPA, ULM, Jl. A.Yani km.36, Banjarbaru, 70714, Indonesia

⁽²⁾Mathematics Study Program, FMIPA, ULM, Jl. A.Yani km.36, Banjarbaru, 70714, Indonesia

*Email: n_salam@ulm.ac.id

Abstract. The agricultural sector is one of the businesses that provides considerable benefits for a country. One of the agricultural sub-sectors that provides benefits is plantations. Coffee plants are familiar plants in the yards of Indonesian people and become one of the leading commodities of the plantation sub-sector. In this case, coffee production is the main target of development in South Kalimantan Province for the next few years. This study aims to obtain the best model so that it can forecast and find out the results of forecasting coffee production in South Kalimantan Province. The Autoregressive Integrated Moving Average (ARIMA) method is the method used in this study to forecast coffee production results in South Kalimantan Province. The results showed that there were 3 ARIMA models that met the forecasting requirements, namely ARIMA (1,2,0) with MSE value 1602502, ARIMA (1,2,1) with MSE value 1049101 and ARIMA model (2,2,1) with MSE value 976202. It can be concluded that the best ARIMA model in forecasting coffee production in South Kalimantan Province uses the ARIMA model (2,2,1) with a smaller MSE value than other models. The forecasting results of coffee production in South Kalimantan Province in 2023 to 2032 tend to decrease, namely with results in 2023 of 718.57 tons, in 2024 of 697.62 tons, in 2025 of 376.87 tons, in 2026 of 344.4 tons, in 2027 of 33.1 tons.

Keywords: *Forecasting, Autoregressive Integrated Moving Average, Coffee Production.*

INTRODUCTION

Indonesia is known as an agrarian country where most of its people work as farmers, a sector that should have become one of the government's top priorities in national development. The agricultural sector is one of the businesses that provide a sizable profit for a country. One of the profitable agricultural sub-sectors is plantations. The plantation sub-sector has plant characteristics that can be grouped into two, namely annual crops and perennial crops. Perennial plants are plants that require a long time to produce, usually an annual production period of up to tens of years, and can be harvested more than once. Examples of annual crops include coconut, oil palm, rubber, cocoa, cloves, coffee, pepper, nutmeg, hazelnut, cloves, cinnamon, vanilla, tea, cottonwood, and so on. While annual crops are plants that can only be harvested once with a life cycle of once a year. Examples of annual crops include sugarcane, citronella, patchouli, and tobacco (Permatasari, 2014)

Coffee plants are plants that are familiar in the yards of the Indonesian people and are one of the leading commodities in the plantation sub-sector. Indonesia is a coffee-producing region with the most specific flavors in the world, according to the International Coffee Organization (ICO), coffee production in Indonesia ranks fourth in the world after Brazil, Vietnam, and Colombia. Coffee production in South Kalimantan Province includes the 4 main plantation commodities, the first is oil palm, the second is rubber, then the third is coconut and the fourth is coffee. According to (Badan Pusat Statistik, 2022) in 2021, the national average of coffee production in South Kalimantan Province is 1.1 thousand tons. There are several aspects why coffee plants are still not evenly distributed across districts/cities in South Kalimantan Province, such as a lack of technical knowledge about good and correct coffee cultivation, then lack of plant maintenance by planters/farmers (low productivity), difficulty getting superior seeds and well, constraints on HR (Human Resources) in the field of coffee plants and the lack of promotion and outreach to the community about the role of coffee in community economic development. Coffee itself is currently in great demand by many people, both young and old, and many young entrepreneurs who are developing their businesses in the coffee sector. In addition, coffee also has other uses, namely as a medicine in the health sector as well as various kinds of preparations, both fresh and pure food and drinks.

In this case, the Province of South Kalimantan has made coffee the main development target for the next few years, therefore there are many ways that can be done, especially by related parties. Especially in this research, it is also explained how to apply statistics to assist related parties in the development of coffee production in South Kalimantan Province, namely the forecasting method. Forecasting is an attempt to predict future conditions through testing past conditions (Handoko, 1984). A time series data is a series of observations

or observations made at a certain time, usually at the same interval (Spiegel, 1994). Autoregressive Integrated Moving Average (ARIMA) is a combination of the moving average method and the autoregressive method, which is a time series data forecasting method that utilizes historical data and current data to produce accurate short-term forecasts. This model does not include independent variables in its formation, this model uses information in the series itself to produce forecasts (Hanke, J. E., & Wichern, 2005). The ARIMA method can also accept all types of data models, although the process must be stationary first. According to (Wei, 2006), the most popular and widely used time series method in forecasting univariate time series data is the ARIMA method.

Based on the discussion above, the ARIMA method is used to predict coffee production in South Kalimantan Province to find out the number of forecasts that will occur. In preparing this method, the first thing that must be done is to make a data plot, and after that to analyze the data so that an appropriate model can be determined for forecasting the amount of coffee production. After getting the best model that has been determined, the forecast for one year or several years to come can be found. With this method, it is hoped that the forecasting data will have a high level of accuracy so that it can be used to predict coffee production in South Kalimantan Province.

MATERIALS AND METHODS

Time Series Analysis

Forecasting is a statistical method that plays an important role in decision-making. One of the methods used in forecasting is the time series method. Future predictions are made based on past information from a variable or past errors. This is called a time series or time series (Makridakis, S., Wheelwright, S. C., & McGee, 1999). The goal of time series analysis is to predict future values with no errors or very small errors. The smaller the deviation or error produced, the better the forecast obtained.

ARIMA method

The forecasting method commonly used is Autoregressive Integrated Moving Average (ARIMA) or Box-Jenkins. Autoregressive Integrated Moving Average (ARIMA) models have been extensively studied by George Box and Gwilym Jenkins (1976), and their name is often symbolized by the ARIMA process applied to time series analysis, forecasting, and control. The ARMA method was developed in three directions, namely identification, efficiency, and assessment procedures (for mixed AR, MA, and ARMA processes), expansion of the results to include seasonal time series, and simple development covering non-stationary processes. (non-stationary process) (ARIMA).

The ARIMA method is grouped into 4 groups, namely the Autoregressive Model (AR), the Moving Average Model (MA) and the Autoregressive Moving Average Model (ARMA) which have the combined characteristics of the AR model and the MA model and the Autoregressive Integrated Moving Average (ARIMA) which involves the process of differencing from ARMA models.

a. Autoregressive Model (AR)

In general, the Autoregressive Model (AR) or ARIMA model (p,0,0) is stated as follows:

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \dots + \phi_p z_{t-p} + e_t \quad (1)$$

Where :

z_t : t-th data

z_{t-p} : data in the t-p period t-p.

p : the p-th autoregressive parameter

e_t : error value at the t-th time.

b. Moving Average Models (MA)

The general Q-order MA process can be written as follows:

$$z_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \dots - \theta_q e_{t-q} \quad (2)$$

Where :

z_t : t-th data

θ_q : moving average parameter with order q

e_{t-q} : error value at t-q

c. Autoregressive Moving Average Model (ARMA)

The AR and MA models can be combined into the same equation, the combination of the two models is called ARMA (p,q) with the general equation being:

$$z_t = \phi_1 z_{t-1} + \dots + \phi_p z_{t-p} + e_t - \theta_1 e_{t-1} - \dots - \theta_q e_{t-q} \quad (3)$$

Where :

p : order AR

q : order MA

Selection of the Best Model

As it is known that no forecasting method can accurately predict the state of the data in the future, therefore every forecasting method will produce errors. If the resulting error rate is smaller, then the forecasting results will be more precise. To determine the best model from several models that meet these requirements, one criterion can be used, namely the Mean Squared Error (MSE)(Durandt, 2019).

Mean Squared Error (MSE) is a criterion for selecting the best model based on the results of the rest of the forecast. The smaller the MSE value means that the estimated value is getting closer to the actual value, or the model chosen is the best model (As'ad, M., Wibowo, S. S., & Sophia, 2017).

$$MSE = \frac{\sum e^2}{n} \quad (5)$$

where :

n : amount of data

e : the difference between the original value and the forecast, with $e = |y - \hat{y}|$.

Research Method

The data used in this study is secondary data obtained from the South Kalimantan Province Plantation and Animal Husbandry Service (Disbunnak). The variable in this study is coffee production level data in South Kalimantan Province in 1998 – 2022 with a total of 25 data in the form of annual data and tons.

This research will be carried out through the following stages:

1. The first step for modeling using the ARIMA method is to identify plot patterns of coffee production data.
2. Conduct a stationary test on coffee production data by looking at the results of the Augmented Dickey-Fuller test, if the data is not stationary, then transformation and differencing are necessary.
3. Doing modeling for the ARIMA model by looking at the ACF and PACF plots.
4. Estimating the appropriate ARIMA (p,d,q) model.
5. Perform a diagnostic check by testing the suitability of the model, namely the white noise test and the normal distribution assumption test of the ARIMA model that has been formed. If the two assumptions are not met, a new model is found that meets these two assumptions.
6. Choose the best model by looking at the minimum Mean Square Error (MSE) value.
7. Forecasting coffee production for the next 5 years with the best model that has been formed.
8. Interpret coffee production with the forecasting model used.
9. Drawing conclusions obtained from the interpretation results.

RESULTS AND DISCUSSION

There are several stages in conducting data analysis using coffee production data in forecasting with the ARIMA method. The stages that will be carried out are:

a. Model Identification

The first step in forecasting using the ARIMA method is to identify the model. Model identification aims to determine the right ARIMA model in forecasting, ARIMA models can be determined if the data is stationary.

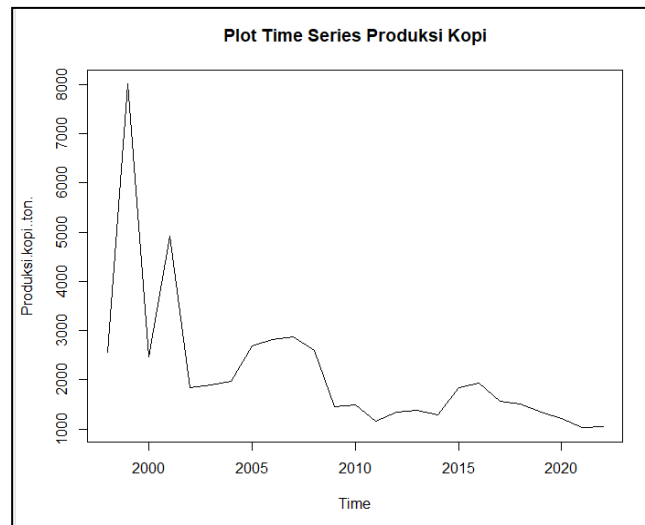


Figure 1. Plot series coffe production

Based on the figure, it shows that coffee production from 1998 to 2022 has decreased which tends not to be constant towards a certain value. Data is said to be stationary if the process does not change with changes in time. The data must be horizontal as long as the data fluctuations are around a constant average value, so it can be concluded that the coffee production data is not stationary in terms of mean or variance. Because graphically the data is not stationary, to deal with data that is not stationary in variance we can perform a Box-Cox transformation. The data can be said to be stationary in terms of variance or variance if the result of the rounded value in the Box-Cox plot is 1.

After doing the calculations to transform the original data by transforming the power of the lambda previously obtained, the results of the Box-Cox transformation are obtained as follows:

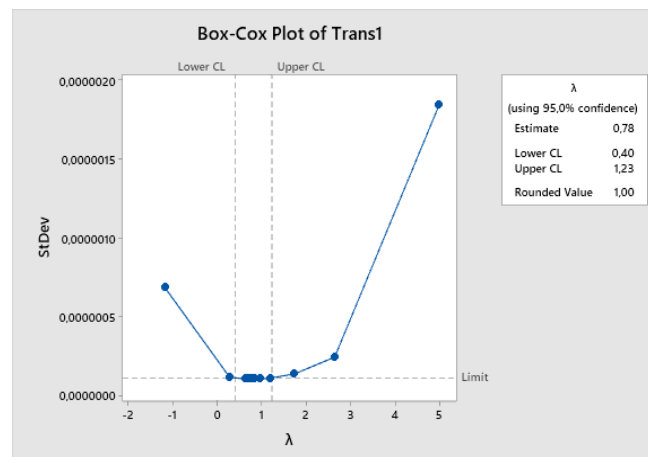


Figure 2. Box-Cox Plot of the Transformations

It can be seen in Figure 2 that the rounded value has a value of 1 and is at the lower limit of 0.40 and the upper limit of 1.23, which means that the data is stationary in variance. So the next step is to perform differencing or differentiation to see whether the data is stationary in the average or not. The plot of the one-time differencing results ($d=1$) as a result of the transformation of coffee production can be seen in Figure 3 as follows:

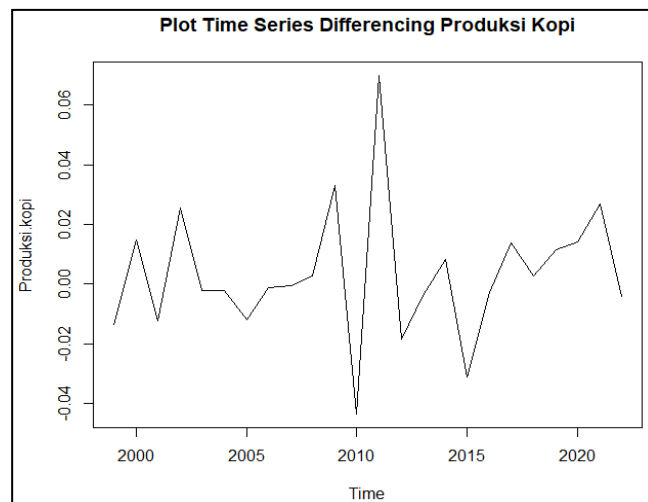


Figure 3. Time Series Plot from differencing 1

After the first differencing was carried out on the coffee production data, it can be seen in Figure 3 above that the data still shows an increase or decrease over time and is not constant within a certain value. This shows that the results of the first differencing data are still not stationary in the average. Because it graphically proves that the data is still not stationary in the mean, an Augmented Dickey-Fuller test is performed from the first differencing data.

Table 1. Augmented Dickey Fuller test results from differencing data 1

Dickey-Fuller	p-value
-2.118	0.5274

From the table, the p-value of 0.5274 is greater than the significance level which still cannot overcome the stationarity of the data in the average, so the next step is to perform a second differencing from the first differencing data. Table of results of differential or differentiation twice ($d = 2$) results of differencing from coffee production can be seen as follows :

Table 2. Augmented Dickey Fuller test results from differencing data 2

Dickey-Fuller	p-value
-3.8517	0.03202

Based on the results of the Augmented Dickey-Fuller test from the second differencing data in Table 2 it shows that the data is stationary in the average. It can be seen from the p-value which is 0.03202 which is smaller than the significance level α which is equal to 0.05. So it can be concluded that the data is said to be stationary in the mean as well as in the variance or variety.

b. Model Estimation

If the data obtained is stationary, the next step is to estimate the model. The differential data ACF and PACF plots are obtained as follows:

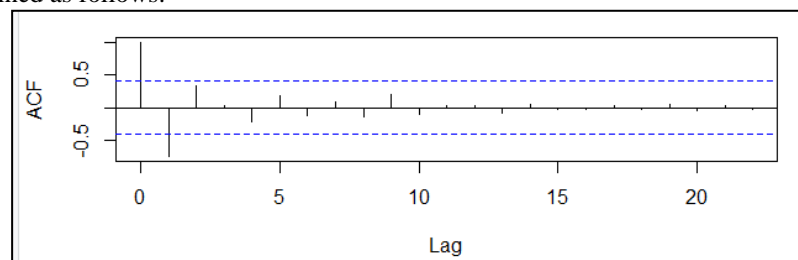


Figure 4. ACF differencing plot 2

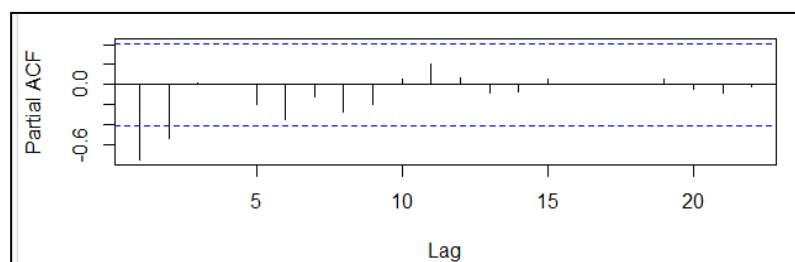


Figure 5. PACF differencing plot 2

Based on Figure 4 and Figure 5, it shows that the data is stationary after differencing. The differencing plots of ACF and PACF show that the data pattern drops exponentially so that it can be said to be cut off. From the results of the analysis, the ACF plot graph is significant at lag 0 and lag 1, then the PACF plot graph is significant at lag 1 and lag 2. It can be said that the data resulting from the second differencing are stationary and a temporary model can be determined, namely with the following details:

Table 3. ARIMA Model Analysis Results

Model	Parameter	Coefficient	Standar Error	MSE
ARIMA (1,2,0)	ϕ_1	-0.9822	0.0252	1602502
ARIMA (1,2,1)	ϕ_1	-0.9489	0.0690	1049101
	θ_1	-0.8588	0.1896	
ARIMA (2,2,0)	ϕ_1	-1.6101	0.1752	1039178
	ϕ_2	-0.6402	0.1822	
	ϕ_1	-0.7556	0.2678	
ARIMA (2,2,1)	ϕ_2	0.2035	0.2730	976202
	θ_1	-0.9771	0.6471	

c. Diagnostic Checking and Selection of the Best Model

To perform diagnostic checking, residual data from each model is required. Then, based on the residual data, several tests and analyzes were carried out. In this case the white noise test and normality test must be fulfilled, to test that the residuals are random can be done with the Ljung-Box test and for normally distributed residuals can be done with the Jarque-Bera test.

Table 4. Results of the Ljung-Box test and the Jarque-Bera test

Model ARIMA	Ljung-Box Test		Jarque-Bera Test	
	<i>x-squared</i>	<i>p-value</i>	<i>x-squared</i>	<i>p-value</i>
(1,2,0)	2.366	0.124	1.8605	0.3945
(1,2,1)	3.493	0.06163	0.64963	0.7227
(2,2,0)	4.6424	0.03119	1.845	0.3975
(2,2,1)	0.79764	0.3718	2.4687	0.291

Based on the table it can be seen that the p-value of the Ljung-Box test of the three ARIMA models, namely ARIMA (1,2,0), ARIMA (1,2,1), and ARIMA (2,2,1) is greater than the significance level or in other words H_0 is rejected, which means that the three ARIMA models whose residuals are formed are random or white noise. While the results of the residual normal distribution test or the Jarque-Bera test can be seen that the p-value of all ARIMA models is greater than the significance level α , which is equal to 0.05, which means that all ARIMA models that are formed have residuals that are normally distributed.

From the things that have been described above, there are only three ARIMA models that are formed which are said to be suitable for use because they meet the requirements for the assumption that the residuals are random and normally distributed. Therefore, the selection of the best model can be seen from the results of model estimation by looking at the smallest error or MSE value. The smaller the MSE value generated, the better the forecasting method used. The following is a summary of the results of the Mean Square Error (MSE) values of the three ARIMA models that are formed which meet the requirements for the assumption that the residuals are random and normally distributed.

Table 5. Summary of Mean Square Error (MSE) values

Model ARIMA	Mean Square Error (MSE)
(1,2,0)	1602502
(1,2,1)	1049101
(2,2,1)	976202

Based on the table described above, of the three models, the ARIMA model (2,2,1) has the smallest Mean Square Error (MSE) value so the best and most appropriate forecasting model to use to forecast coffee production in South Kalimantan Province is the ARIMA(2,2,1). The ARIMA model (2,2,1) can also be seen with the following equation:

$$z_t = 1.2444 z_{t-1} + 2.7147 z_{t-2} + 0.3486 z_{t-3} - 0.2035 z_{t-4} + 0.9771 e_{t-1} + e_t$$

d. Forecasting

The following is the result of forecasting coffee production for the next 5 years.

Table 6. Forecasting Results

Periods	Forecasting
2023	718.57
2024	697.62
2025	376.87
2026	344.4
2027	33.1

Based on the forecasting results in the figure above, it can be seen that the forecasting results for coffee production data have decreased significantly, this could have happened because modeling and forecasting by the ARIMA model are mostly based on the most recent historical data. When there is an increase or decrease in the data, it is natural for the predicted value to depend on the latest fluctuating data (Hutasuhut, A. H., Anggraeni, W., & Tyasnurita, 2014).

CONCLUSION

Based on the research results in the discussion, it can be concluded that the best model of the ARIMA method used to forecast coffee production in South Kalimantan Province is the ARIMA model (2,2,1). This is indicated by the smaller MSE value compared to the other models. The ARIMA model (2,2,1) can also be seen with the following equation:

$$z_t = 1.2444 z_{t-1} + 2.7147 z_{t-2} + 0.3486 z_{t-3} - 0.2035 z_{t-4} + 0.9771 e_{t-1} + e_t$$

From the best-selected model, the results of forecasting coffee production in South Kalimantan Province for 2023 to 2027 are obtained, namely:

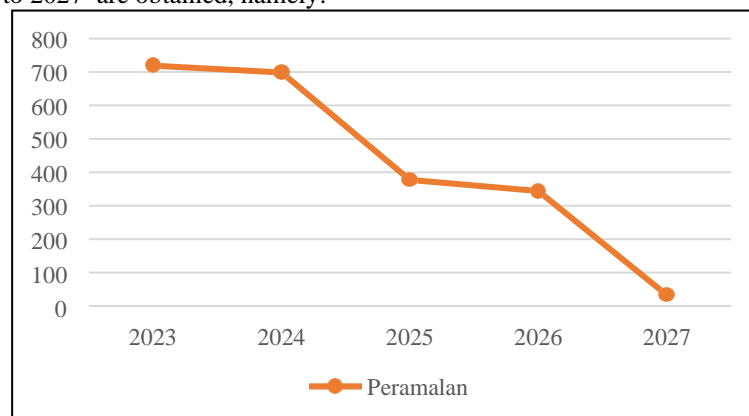


Figure 6. Forecasting Results of Coffee Production in 2023 – 2027.

Forecasting results of coffee production data tend to decrease, this could happen because modeling and forecasting by the ARIMA model are mostly based on the most recent historical data. When there is an increase or decrease in the data, it is only natural that the predicted value depends on the latest fluctuating data.

REFERENCES

- [1] As'ad, M., Wibowo, S. S., & Sophia, E. (2017). Peramalan Jumlah Mahasiswa Baru Dengan Model Autoregressive Integrated Moving Average (Arima). *J I M P - Jurnal Informatika Merdeka Pasuruan*, 2(3), 2(3), 20–33. <https://doi.org/10.37438/jimp.v2i3.77>.
- [2] Badan Pusat Statistik. (2022). *Statistik Indonesia 2022*. In D. D. Statistik (Ed.), *Statistik Indonesia 2022*.
- [3] Durandt, H. R. B. (2019). Model Peramalan Laka Lantas Polres Kolaka menggunakan Metode Autoregressive Integrated Moving Average (ARIMA). *Universitas Islam Negeri Alauddin Makassar*.
- [4] Handoko, T. H. (1984). *Dasar-Dasar Manajemen Produksi Dan Operasi*. BPFE.
- [5] Hanke, J. E., & Wichern, D. W. (2005). *Business Forecasting* (8th ed.). Pearson Prentice Hall, New Jersey.
- [6] Hutasuhut, A. H., Anggraeni, W., & Tyasnurita, R. (2014). (2014). Pembuatan Aplikasi Pendukung Keputusan Untuk Peramalan Persediaan Bahan Baku Produksi Plastik Blowing dan Inject Menggunakan Metode ARIMA (Autoregressive Integrated Moving Average) di CV. Asia. *Jurnal Teknik POMITS*, 3(2), A-169.
- [7] Makridakis, S., Wheelwright, S. C., & McGee, V. E. (1999). *Metode dan Aplikasi Peramalan* (2nd.ed). Erlangga, Jakarta.
- [8] Permatasari, D. (2014). *Analisis Pendapatan Usahatani Gula Tumbu (Kasus Kecamatan Dawe Kabupaten Kudus)*. Universitas Diponegoro Semarang.
- [9] Spiegel, M. R. (1994). *Teori dan Soal-Soal Statistika* (2nd ed). Erlangga.
- [10] Wei, W. W. S. (2006). *Time Series Analysis: Univariate and Multivariate Methods* (Second Edi). Pearson Education, United States of America. <https://doi.org/10.2307/1269015>.

Predicting the Consumer Price Index in South Kalimantan Province Using Support Vector Regression Method

Rahma Dina Nur Azizah^{*(1)}, Dewi Sri Susanti⁽¹⁾, and Selvi Annisa⁽¹⁾

⁽¹⁾*Department of Statistics, Faculty, Universitas Lambung Mangkurat*

**Corresponding author: 19111017220004@mhs.ulm.ac.id*

Abstract. Inflation, which is a phenomenon that represents an increase in the price of goods and can cause a country's economic slump, is one of Indonesia's economic concerns. The Consumer Price Index (CPI) is a popular indicator for measuring inflation since it allows food price stability to be controlled by knowing the CPI number ahead of time. Support Vector Regression (SVR), a prediction method that finds the optimum hyperplane approach in the regression function that gives the smallest error, is one of the statistical methods that may be used to predict CPI. This study employs CPI data for the food ingredients category in the city of Banjarmasin from 2014 to 2022. For the parameter value $C = 1$ and $\epsilon = 0.0001$, the best predictive model is obtained using the SVR method with the RBF kernel. The prediction model developed is very good and practicable, based on the MAPE value and the coefficient of determination acquired from the prediction. The CPI projection results in Banjarmasin for the foodstuffs category for the next 12 months show an increase, allowing the government and affiliated parties to make policy or anticipate steps.

Keywords: Consumer Price Index, Support Vector Regression, Banjarmasin.

INTRODUCTION

One of the economic problems in Indonesia is inflation, which is a condition that reflects an increase in the price of goods because it can cause a country's economic downturn (Aris Kencono, 2018). Factors that affect a country's inflation rate are usually influenced by domestic conditions and the indicator that is often used to measure the inflation rate is the Consumer Price Index (CPI) (Hadi et al., 2019). The CPI is the average change in the price of commodity packages consumed by residents/households within a certain period of time and is measured based on the Cost of Living Survey. Changes in the CPI from time to time are an important indicator to describe the rate of inflation/deflation (Priyono et al., 2016).

CPI by expenditure group as a whole is classified into 11 groups and 43 sub-groups. The foodstuffs group consists of several supporting sub-groups, namely grains, tubers, meat, fresh fish, preserved fish, eggs and milk, vegetables, nuts, fruits, spices, fats and oil, and foodstuffs. Foodstuff commodities have an important role because they are consumed every day and become the main staple that supports people's lives. It is necessary to maintain a stable level of food commodity prices in the community in the current and future periods, bearing in mind that the stability of food commodity prices is an indicator of economic progress in a region. The government as a policy maker needs to realize sustainable economic growth so that the socio-economic conditions of the community become more stable. Price stability can be controlled by making predictions using the Support Vector Regression (SVR) method. SVR is a method that seeks the best hyperplane in the form of a regression function by minimizing the error as much as possible while maximizing the margin. SVR is capable of handling both linear and non-linear data.

MATERIALS AND METHODS

Materials

The data used in this research is secondary data Consumer Price Index Expenditure Group, Foodstuffs Sub-Group. The total data is 108 observational data from the period January 2014 – December 2022 (BPS, 2022). The independent variable (X) in this study is CPI data for the Foodstuffs Group one month previously which is used to predict CPI data for the following month and it is assumed that this month's CPI is influenced by the previous month. The determination of the CPI data variable for the city of Banjarmasin is determined based on the time lag obtained from the PACF plot.

Descriptive Analysis

According to Gresik (2018), descriptive analysis is part of statistics that studies how to collect data and present data so that it is easy to understand. Quantitative descriptive analysis is a statistical technique used to describe and explain the characteristics of numerical data. Numerical data can also be translated into tables and graphs such as histograms, frequency polygons, and box plots to assist in understanding the characteristics of the data, the goal is to understand the distribution, pattern, and distribution of data through descriptive statistics such as mean, standard deviation, and frequency so the data can be presented in a form that is easier to understand and read.

Time Series Analysis

According to Box and Jenkins (2015) a time series is a group of observational values obtained at different points in time with the same time interval and the data sequences are assumed to be interconnected with each other, so a time series model is a model in which observations are one with the other. the others are interconnected. The aspect that needs to be considered in analyzing time series data is to look at the pattern of the data studied. There are four kinds of data patterns in the prediction of time series data, namely, horizontal/constant data patterns, seasonal/seasonal data patterns, trend data patterns and cyclical data patterns. Horizontal data patterns occur when the data fluctuates or goes up and down around a constant average value, seasonal patterns occur when a series is influenced by seasonal factors. Trend patterns occur when there is a long-term increase or decrease in data and cyclical patterns occur when the data is affected by long-term economic fluctuations (Lopez-Martin et al., 2018).

Data Normalization

Data normalization is a process that aims to equalize the scale or range of values that exist between different variables so that data can be compared and analyzed more accurately. While data denormalization is the process of returning normalized values to actual data. According to (Dewi et al., 2018) the standard scale that is usually used ranges from 0 to 1. The equations used for normalizing and denormalizing data are shown in the following equation

$$x' = \frac{(x - x_{min})}{x_{max} - x_{min}} \quad (1)$$

$$x = x_{min} + (x'(x_{max} - x_{min})) \quad (2)$$

- x' : Data normalization results
- x : The data value to be normalized
- x_{max} : The maximum value of the data
- x_{min} : The minimum value of the data

Support Vector Regression

The Support Vector Regression (SVR) discovered by Vladimir N. Vapnik (1999) is the result of a modification of the Support Vector Machine (SVM) method used to solve prediction problems (Smola & Schölkopf, 2004). The concept of the SVR algorithm can produce good forecasting values because SVR has the ability to solve overfitting problems. Overfitting is the behavior of the data during the training phase resulting in almost perfect prediction accuracy. The goal of the SVR algorithm is to find the dividing line or it can be called the best hyperplane. The best hyperplane can be found by measuring the margins with that hyperplane. While the margin is the distance from the hyperplane to the closest data. The data closest to the margin is called the support vector. It can be concluded that SVR is a method by obtaining the best hyperplane in the form of a regression function by making the smallest possible error by maximizing the margin. SVR illustration can be seen in the following figure.

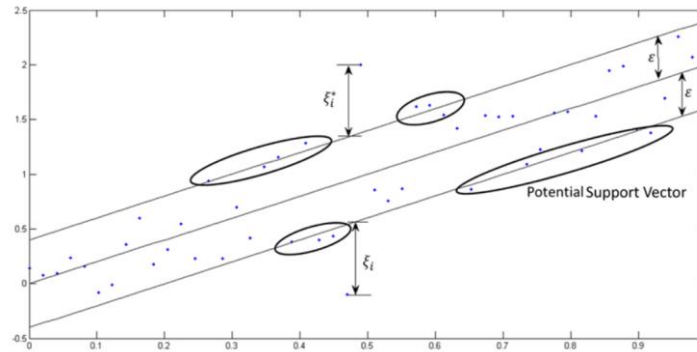


Figure 1. SVR illustration

The figure shows a line in the middle or a hyperplane with two dividing lines, namely the epsilon boundary $\varepsilon (+)$ and the epsilon boundary $\varepsilon (-)$, while some of the circled points (data points) are potential support vectors. These points are potential boundaries, so they can enter a zone while minimizing the *epsilon* ε as much as possible. Points that are outside the ε range can be added to the slack variable ξ .

The loss function and kernel function play an important role in SVR. Prediction with the SVR method is suitable for high-dimensional data sets and can be used for cases of non-linear linear data using kernel functions (Dewi et al., 2018). The loss function is used to determine how much error is acceptable in predicting. This loss function allows the model to limit the impact of outliers and strengthen consistency in predicting targets (Lopez-Martin et. al., 2018). Meanwhile, the kernel function is to transform data in a non-linear input space into a high-dimensional feature space to find the optimal hyperplane. Several types of kernels commonly used in SVR are linear, polynomial, radial basis function (RBF).

1. Linear Kernel

$$K(x_i, x_j) = x_i^T x_j \quad (3)$$

2. Radial Basis Function Kernel

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2) \quad (4)$$

3. Polynomial Kernel

$$K(x_i, x_j) = (\gamma x_i^T x_j + 1)^d \quad (5)$$

Grid Search Optimization

Grid Search Optimization is a method for finding the best hyperparameters in a Machine Learning model by setting and evaluating selected parameter sets in a grid or table which involves combining all the parameters in the grid and training the model for each combination of parameters in order to obtain the model with the best performance. Cross validation is used to evaluate the performance of the model for each combination of these hyperparameters. The results of this cross validation are used as a basis for selecting the best hyperparameters in the grid search (Intan Ari Budiastuti, 2017).

In cross validation, the training data is divided into several k folds that are side by side and equal in size, each part is used once as validation data and the rest is used as training data. Each model of the hyperparameter combination is then calculated the average MSE score for all folds (Eka Patriya, 2020). In K-fold cross validation, the data is divided randomly into k parts so that the time sequence in the data is not considered, but in some cases, the time sequence in the data has an influence on the prediction results of the model (Raschka, 2015). In time series data, the cross validation technique that is often used is Time Series Split. Time Series Split is to divide the time series data into several successive subsets and carry out the validation process for each subset.

Model Accuracy

MAPE (Mean Absolute Percentage Error) is one of the methods used to measure the ability of the predictive model to predict the target variable. MAPE measures the absolute error in percentage of the actual target (Montgomery, 2008). The MAPE equation is shown in the following equation.

$$MAPE = \frac{\sum_{i=1}^n \left| \frac{x_i - \hat{f}_i}{x_i} \right|}{n} \times 100\% \quad (6)$$

x_i : actual data for period i
 f_i : the predicted value of the i-period
 n : number of data

Research procedure

1. Collect data available on the BPS website, namely CPI data for the Foodstuffs Group in the city of Banjarmasin for the period January 2014 - December 2022
2. Conduct a descriptive analysis to see the distribution of data and describe the CPI data for the Foodstuffs Group in the city of Banjarmasin using a line chart to see patterns of data changes
3. Data normalization by determining the dependent variable (Y) and independent variable (X) using PACF plots
4. Dividing the dataset into training data by 80% and data testing by 20%. 80% of the training data is used to build and ensure the model has enough data to study patterns in the data so that it can make good predictions on the training data. Meanwhile, 20% of the data testing is used to evaluate the performance of the model by comparing the predicted results with the actual values in the testing data. This helps ensure that the model is not overfitting the training data and has the ability to accurately predict results on new data
5. Determine the type of kernel function used for forecasting, namely linear kernel, polynomial kernel and radial basis function (RBF) kernel
6. Perform parameter tuning to find the most optimal parameter values that produce the smallest MSE values for the parameters owned by each kernel using the grid search optimization algorithm with time series split cross validation
7. Establish a CPI data SVR prediction model with optimal parameters
8. Determine the best predictive model by evaluating the best accuracy value by looking at the MAPE value
9. Make predictions for the CPI of the Foodstuffs Group in the city of Banjarmasin for the next 12 months.

RESULTS AND DISCUSSION

Descriptive Analysis

Descriptive analysis in this study was used to provide an overview of the CPI data for the foodstuffs group in South Kalimantan Province from January 2014 to December 2022. Figure 4.1 presents a descriptive analysis of the CPI data:

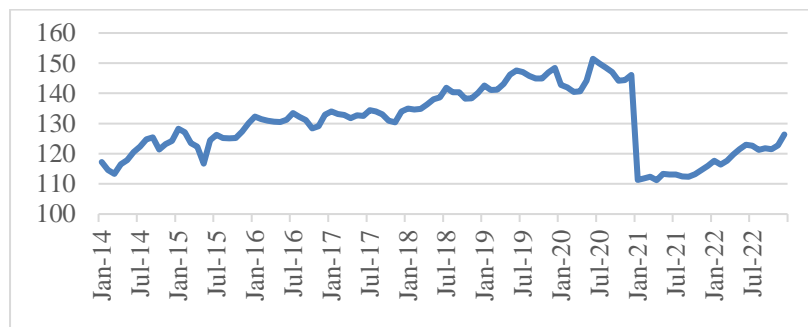


Figure 2. CPI Foodstuffs Group, City of Banjarmasin

Table 1. Statistics Descriptive

Data	N	Mean	Standard Deviation	Min	Max
CPI Foodstuffs Group, City of Banjarmasin	108	130,15	11,02	111,22	151,46

Based on Figure 2, it is known that the movement of the CPI in the city of Banjarmasin from January 2014 to December 2022 was quite volatile, so that at the beginning of 2020 the CPI dropped significantly due to the impact of the Covid-19 pandemic which resulted in lower prices for food commodity packages. The CPI for the materials group in the city of Banjarmasin for 7 years was the lowest at 111.22 which occurred in early 2021 when the pandemic was still occurring at that time and the highest CPI was 151.46 which occurred in mid-2020. The following presents descriptive statistics from the CPI group Banjarmasin food ingredients.

Data Normalization

Data normalization is carried out so that the resulting SVR model has a good level of accuracy. In this study, using the PACF plot to determine the lag that affects the data, then dividing the data into training data and testing data.

Table 2. Data Normalize of Banjarmasin's CPI

X	Y	X'	Y'
117,2	114,59	0,148608	0,083748
114,59	113,32	0,083748	0,052187
⋮	⋮	⋮	⋮
121,77	121,5	0,262177	0,255467
121,5	122,73	0,255467	0,286034

Determination of Influenced Lags Using PACF Plots

CPI data is time series data whose observed values are influenced by previous observed values (lag). Determination of influential lags can be identified using a Partial Autocorrelation Function (PACF) plot by identifying lags that fall outside the significant boundary line. This significant lag will be used to determine the dependent variable (Y) and the independent variable (X). Figure 4 is the result of the PACF plot on CPI data for the city of Banjarmasin.

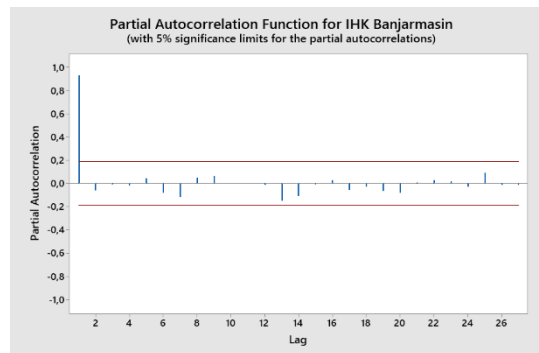


Figure 3. PACF Plot of Banjarmasin's CPI Data

In the figure above, it can be seen that the significant lag is shown by the blue line (partial autocorrelation value) that intersects the confidence interval (red straight line). In the CPI data in the city of Banjarmasin, there is a significant lag based on the PACF plot, namely lag 1 so that in this study the independent variable (X) used as input is x_{t-1} . CPI data for the Banjarmasin city food ingredients group will be divided into training data (80%) and testing data (20%). Complete distribution of data can be seen in the following Table 3.

Table 3. Data Splitting

IHK City	Data Splitting				
	Training 80%	Periode	Testing 20%	Periode	Number of Data

Tuning Parameters

The best parameter will be obtained from the best MSE value for each Linear, RBF, and Polynomial kernel shown in Table 4.

Table 4. Tuning Parameters

Kernel	Best Parameter	Parameters Value	MSE
Kernel	C	10	0,011502
Linear	ϵ	0,01	
Kernel RBF	C	1	0,011357
	ϵ	0,0001	
	γ	1	
Kernel	C	1	0,0184
Polinomial	ϵ	0,001	
	γ	1	
	d (degree)	2	

Based on the values shown in Table 4, the best parameters for the Linear kernel are $C = 10$ and $\epsilon = 0,01$ with the best MSE value in the Linear kernel of 00115. In the RBF kernel, the best parameters produced are $C = 1$, $\epsilon = 0,0001$, and γ of 1 with the best MSE value in the RBF kernel of 0,011357. In the Polynomial kernel, the best parameters produced are $C = 1$, $\epsilon = 0,001$, $\gamma = 1$, and $d = 2$ with the best MSE value in the Polynomial kernel of 0,0184.

The best parameter in the Polynomial kernel gives a greater MSE value than the Linear kernel and the RBF kernel. As for the Linear kernel and the RBF kernel, the MSE value of the RBF kernel is smaller than that of the Linear kernel. Therefore, the RBF kernel will be used to create a CPI prediction model for the city of Banjarmasin because it produces the smallest MSE value among the three kernels used.

Determination of the Best SVR Model and Model Evaluation

Visualization of the SVR during model formation using training data and model testing using data testing with the RBF kernel on CPI data for the city of Banjarmasin can be seen in Figure 4.

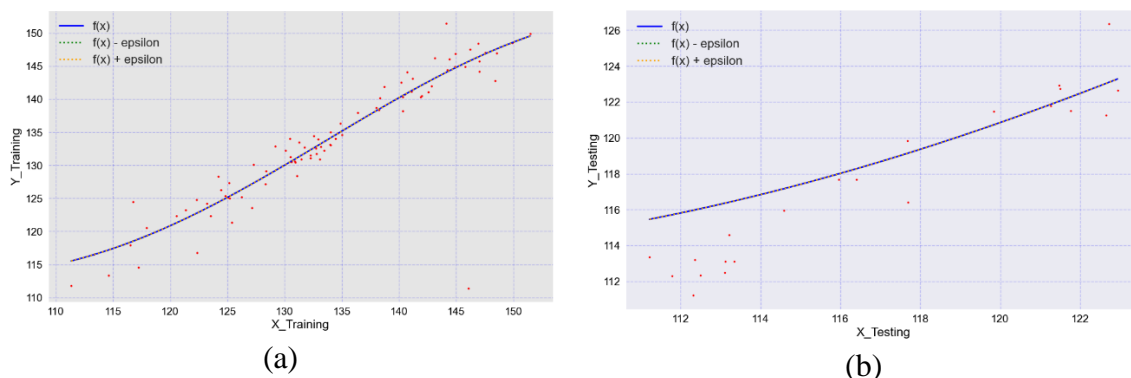


Figure 4. Visualization of the Establishment of the SVR CPI Model for the City of Banjarmasin

Based on Figure 4 it can be seen that the data points for the CPI in the city of Banjarmasin have a non-linear distribution following the line. From the two figures, it can be seen that the hyperplane $f(x)$ almost coincides with the upper and lower bounds or $f(x) \pm \epsilon$ because the optimal ϵ value uses the RBF kernel at the Banjarmasin CPI of 0,0001.

The evaluation of the Banjarmasin CPI is shown by the accuracy of the coefficient of determination R^2 and the MAPE value will be described in Table 7 below.

Table 5. Model Evaluation

Kernel	R^2	MAPE
RBF	0,9048	0,9492

Based on the accuracy value from Table 5, the accuracy of the best R^2 and MAPE in the Banjarmasin CPI using data testing with a data division of 80%:20% produces an $R^2 = 0,9048$ and a MAPE value of 0,9492%. This shows that the comparison between the CPI prediction results for the city of Tanjung on testing data and the actual value of the dependent data testing data using parameters $C = 1, \epsilon = 0,001$ and $\gamma = 1$ can be said to be very good. The following shows a comparison graph of the predicted value of the actual CPI data for the city of Banjarmasin on data testing.

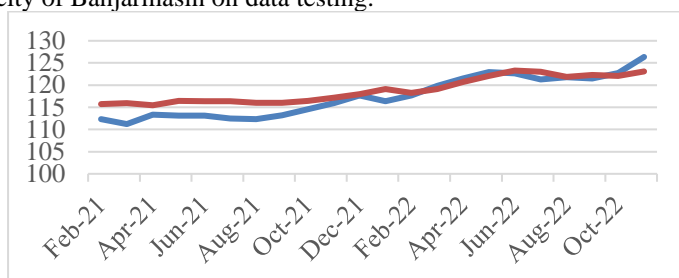


Figure 5. Graph of Comparison of Actual and Predicted CPI Values for the City of Banjarmasin on Testing Data

Based on Figure 5, a comparison of the predicted results with the actual value of the Banjarmasin CPI using the RBF kernel is obtained. The horizontal axis shows the period in the testing data and the vertical axis shows the CPI value in the testing data. The blue line is the actual data, while the red line is the value of the predicted data. On the graph, the plot of the resulting predicted values almost coincides with the actual data, which means that the CPI predictions are not much different so that the SVR model built is feasible to use to predict the Banjarmasin City CPI.

Table 6. Comparison of Actual and Predicted Values in CPI Testing Data for the City of Banjarmasin

Period	Actual Data	Predict Data	<i>error</i>
Feb-21	112,31	115,71	3,401
Mar-21	111,22	115,96	4,740
Apr-21	113,34	115,46	2,120
⋮	⋮	⋮	⋮
Sep-22	121,5	122,28	0,782
Oct-22	122,73	122,06	0,673
Nov-22	126,35	123,10	3,250

Prediction Results for The Next 12 Months

Predictions will be made for the next 12 months with the assumption that this month's CPI is influenced by the previous month's CPI. The following table shows the CPI prediction for the city of Banjarmasin for the next 12 months.

Table 7. Prediction Results for The Next 12 Months

Period	Banjarmasin	Period	Banjarmasin
Jan-23	123,58	Jul-23	124,83
Feb-23	123,85	Aug-23	124,98
Mar-23	124,08	Sep-23	125,11

Apr-23	124,3	Oct-23	125,24
Mei-23	124,49	Nov-23	125,36
Jun-23	124,67	Des-23	125,47

CONCLUSION

The best prediction model was formed using the SVR method using the RBF kernel on CPI data for the food ingredients group in the city of Banjarmasin. Each parameter that produces the best prediction, namely, the city of Banjarmasin, using parameters $C = 1$, $\varepsilon = 0,0001$ and $\gamma = 1$ produces a MAPE value of 0,9492% and $R^2 = 0,9048$. Based on the MAPE value and the coefficient of determination R^2 obtained from the predictions, it can be concluded that the prediction model formed is very good and feasible. The prediction results for the CPI for the foodstuffs group in the city of Banjarmasin for the next 12 months show a tendency to increase each month, so that the government and related parties can make a policy or anticipate it.

ACKNOWLEDGEMENT

First and foremost, i would like to thank our supervisor, Mrs. Dewi Sri Susanti, S.Si., M.Si and Mrs. Selvi Annisa S.Si., M.Si for their invaluable guidance and support throughout the research process. Their expertise and insights were instrumental in shaping the direction and focus of our research. We are also grateful to the Department of Statistics at Lambung Mangkurat University for providing us with the resources and support we needed to complete this project

REFERENCES

- [1] Amanda, R., Yasin, H., & Prahutama, A. (2014). Analisis *Support Vector Regression* (SVR) Dalam Memprediksi Kurs Rupiah Terhadap Dollar Amerika Serikat. *Jurnal Gaussian*, 3(4), 849–858. <https://doi.org/10.14710/J.GAUSS.3.4.849-858>
- [2] Aris Kencono. (2018). Analisis inflasi IHK dan Inflasi IHP di Indonesia (Periode 2000:T1 – 2016:T4). Skripsi. Lampung: Universitas Lampung.
- [3] Asyiva, A. (2019). Prediksi Laju Inflasi Indonesia Menggunakan Metode *Support Vector Regression* Dengan *Kernel Radial Basis Function*. <https://repository.uinjkt.ac.id/dspace/handle/123456789/47435>
- [4] Awad, M., & Khanna, R. (2015). *Support Vector Regression*. 67–80. https://doi.org/doi:10.1007/978-1-4302-5990-9_4 10.100
- [5] Berrar, D. (2018). Cross Validation. In *Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioinformatics* (Vols 1–3, pp. 542–545). Elsevier. <https://doi.org/10.1016/B978-0-12-809633-8.20349-X>
- [6] Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time Series Analysis: Forecasting and Control, 5th Edition* / Wiley.
- [7] Chang, P. C., Wang, Y. W., & Liu, C. H. (2007). The development of a weighted evolving fuzzy neural network for PCB sales forecasting. *Expert Systems with Applications*, 32(1), 86–96. <https://doi.org/10.1016/J.ESWA.2005.11.021>

Intervention Analysis in Forecasting the Number of Arrival Passangers at Syamsudin Noor Banjarbaru

Ratna Annisa Kumala Dewi^{*(1)}, Nur Salam⁽¹⁾, and Yeni Rahkmawati⁽¹⁾

⁽¹⁾Department of Statistics, Faculty of Mathematics and Natural Sciences, Lambung Mangkurat University

*Corresponding author: ratnaannisa07@gmail.com

Abstract. Air transportation as a mode of transportation that supports population mobility has an important role in increasing regional growth and development, catalyzing industrial, tourism and trade activities. Increases and decreases in passengers can be caused by certain events such as national disasters. The Covid-19 pandemic caused a significant decrease in the number of arrival passengers at Syamsudin Noor airport. This study aims to predict the number of passengers arriving at Syamsudin Noor airport when the intervention occurs. By using data on the number of passengers arriving at Syamsudin Noor airport from January 2011 - March 2023. The results showed that the SARIMA (0,1,1) (1,0,0)¹² model with intervention order (0,4,1) is the best model for forecasting because it has the smallest AIC value. The resulting intervention modeling shows that the impact of the intervention in April - July 2023 tends to increase from the previous period. Therefore, the South Kalimantan government can make efforts that can boost the growth of air transportation users, such as developing tourism potential and continuing to carry out promotions in order to attract tourists to stay in South Kalimantan so that they can help increase economic growth after the Covid-19 pandemic.

Keywords: Intervention Analysis, Covid-19 pandemic, The Number of Arriving Passengers, Syamsudin Noor International Airport.

INTRODUCTION

Air transportation as one of the mode of transportation that supports population mobility plays a key role in increasing growth and development of regions, a catalyst for industrial, tourism and trade activities (Adisasmita, 2011). Increases and decreases in passenger air may be caused by religious holidays, holidays and certain events such as national disasters. On March 2, 2020, Coronavirus Disease 2019 (Covid-19) was first found in Indonesia (Sukur et al., 2020). The Covid-19 pandemic was a national disaster that caused the decline in air transport passengers. In order to prevent the spread of the virus, the government is calling on Indonesian people to reduce their mobility and activities outside the home. One of the airports that has been reduced in passenger numbers as a result of government policies regarding the Covid-19 pandemic case is Syamsudin Noor airport. Based on data from Badan Pusat Statistik (BPS) of South Borneo (Figure 1), compared to February 2020. In March 2020, the number of passengers has risen to 101.822, with a decrease in sales from February 2020 at 131.822.

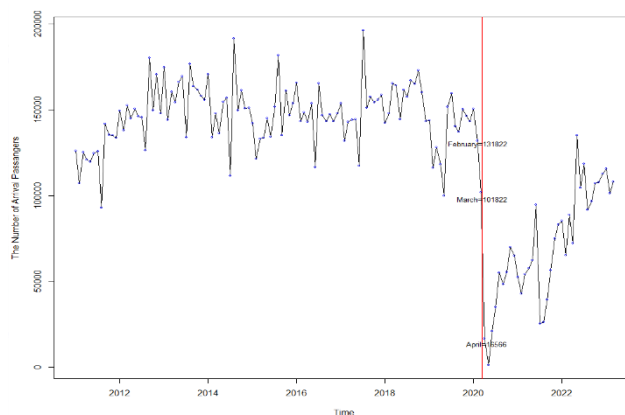


Figure 1. Data plot the number of passengers arriving from January 2011- March 2023 (source: central statistical office of south kalimantan 2011-2023, treated)

According to (Kamaluddin, 2003) flights are able to grow a variety of economic activities, for from flight activities passengers return not only from the countryside to the country where they come from but also to tourists or tourists. According to data from BPS, economic growth in south kalimantan fell to 1.82. Therefore, the decline and growth in the number of passengers arriving at the Syamsudin Noor airport should be forecast, so that governments can prepare development plans specifically for transport and tourism and economic recovery. According (Box et al, 2016) the time series data which changes in patterns significantly and is affected by events known to exist can be modeled with an intervention model, as an intervention model can overcome extreme patterns following previous data patterns. Data on the number of passengers arriving at the Syamsudin Noor of January 2011 - March 2023 saw an intervention in the Covid-19 pandemic that occurred in March 2020. Based on this, the study was determined to predict the number of passengers arriving at the Syamsudin Noor airport, using analysis methods for intervention.

MATERIALS AND METHODS

Materials

The data used is the data of the number of passengers arriving at the Syamsudin Noor airport which is monthly time data in a year with a variable number of passengers arriving. There are 147 observations from January 2011 to March 2023 that sourced to Badan Pusat Statistik of South Borneo provincial office page and can be accessed through <https://kalsel.bps.go.id>.

Time Series Analysis

The time line is a set of observations taken periodically at regular intervals (Box et al., 2016). According to Kitagawa (2010) one of the purposes of the time line analysis is to predict future behavior by using various information from current and past observations.

There are four kinds of data pattern time rows (Hanke & Winchern, 2014) which are horizontal patterns in which data patterns are around constant middle value, trends in data patterns that increase or decrease over a long period of time, seasonal patterns that if data have the same pattern over time over the same period over time because of seasonal factors, And the cyclical pattern is that there are fluctuations around the trend lines that are usually caused by long-term economic factors.

Data Stationary

A data is said to be stationary if data fluctuations are around an average value (mean) and its variance is constant all the time (Frechting, 2011). The stationary in mean can be seen from the Autocorrelation Function (ACF) plot and the Dickey-Fuller test, if it is not stationary in the rate one by one to do differencing. Whereas diversification is detected by the use of Box-Cox plot (Box et al., 2016). Data is arguably a stationary in great variation if the lambda value approaches the one or the staggering value of the plot of Box-Cox is worth one. If it is not stationary in diversity, it can be overcome by Box-Cox transformational methods.

Seasonal Autoregressive Integrated Moving Average (SARIMA) Model

The model SARIMA has a common notation SARIMA $(p, d, q)(P, D, Q)^s$ which is defined as follows (Box et al., 2016):

$$\phi_p(B)\Phi_p(B^s)(1-B)^d(1-B^s)^D Z_t = \theta_q(B)\Theta_q(B^s)a_t \quad (1)$$

With Z_t is observations at the t time, ϕ_p represents AR parameters (p), Φ_p are seasonal AR parameters (P), θ_q are MA parameters (q), Θ_q are seasonal MA (Q) parameters and a_t are residuals.

1. SARIMA Model Identification

SARIMA model identification is made by establishing seasonal and non-seasonal orders. The approval of the SARIMA model order is done by seeing ACF and PACF plots from data that have been stationary (Hanke & Winchern, 2014) ACF plots and PACF has two possible patterns, which are dropped dramatically (cuts off) after the lag (p,q, p,q) and decline slowly (tails off).

2. Significance test of the parameter may be made using t-test (Agresti, 2018).

3. The test of autocorrelation can be done by formal test using Ljung-Box test statistics. Where the rest can be said to be free when p-value $> \alpha$.

4. The normality test could be done by using the Jarque-Bera test or by using the Kolmogorov-Smirnov test (Daniel, 1989).

Intervention Model

The model of intervention is a combination of the SARIMA model and the function of a free shapeshifter with the following general figure (Box et al., 2016):

$$Y_t = f(\beta, I_t) + \frac{\theta_q(B)\theta_Q(B^S)}{\phi_p(B)\phi_P(B^S)(1-B)^d(1-B^S)^D} a_t \quad (2)$$

where, $\frac{\theta_q(B)\theta_Q(B^S)}{\phi_p(B)\phi_P(B^S)(1-B)^d(1-B^S)^D} a_t$ was a SARIMA model before the intervention. The magnitude and length of the impact of interventions on data rows of time is defined by the deformity function ($f(\beta, I_t)$) generally defined as follows:

$$f(\beta, I_t) = \frac{\omega_s(B)}{\delta_r(B)} B^b I_t \quad (3)$$

with $\omega_s(B)$ form the parameters of the order s, $\delta_r(B)$ constitutes the parameters of the order r and I_t constitutes the intervening function. Based on time of occurrence, there are two features of the intervention model, one with the function of the pulse input and step. The only interventions that only exist at a given time are called the just function of the pulse. The toggle of the pulse function is suggested as follows:

$$I_t = P_t^{(T)} = \begin{cases} 0, & t \neq T \\ 1, & t = T \end{cases} \quad (4)$$

The step function interventions are interventions that occur over long periods of time. The toggle of intervention function in step is chosen as follows:

$$I_t = S_t^{(T)} = \begin{cases} 0, & t < T \\ 1, & t \geq T \end{cases} \quad (5)$$

The order b,s,r is an important order in an intervention model that can be identified using the SARIMA print index of data before the intervention. The order b,s,r can be identified by the following conditions (Aritara, 2011):

- The b order represents the number of time periods before the effects of intervention affect the data sequence of time. The b order is worth zero, if the impact of a direct intervention is felt at the time of an intervention. The order b is worth one, if the impact of an intervention is felt one period after the intervention, and so on.
- The order of s represents much of the time period it takes for the data to be calculated back stable from the time of the intervention.
- The order r is the next lag after the b and s. the order r as seen from no pattern on the parenthetical graph. The order r is worth 0 when it does not form a clear pattern, the order r is worth 1 when it forms an exponential pattern, the order r will be worth 2 when it forms a sinusoidal pattern.

Choosing the Best Model

There are some criteria that can be used in selecting the best models of some possible models, one of which is akaike's information criterion (AIC). AIC has the following equation (Mantgomery, 2015):

$$AIC = \ln \frac{(\sum_{t=1}^n e_t^2)}{n} + \frac{2h}{n} \quad (6)$$

with h is the sum total of model parameters and e_t is the sum of square model.

The best model selection criteria based on the correct model in forecasting can use the Root Mean Square Error (RMSE) value. Formula RMSE is as follows (Mantgomery, 2015):

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (F_t - \hat{F}_t)}{n}} \quad (7)$$

Research Procedure

The research will be conducted through the following steps which is, first make plot of time series data the number of passengers arriving from January 2023- March 2023 to look at the data characteristics and patterns and determine the appropriate analysis models, and then divide the data into two parts the data before the intervention (January 2011 – February 2020) and the data during the intervention (March 2020 - March 2023). After that, build SARIMA model for data before intervention with the step is:

1. Checking data stationary in mean and varian.
2. Identifying tentative models based on ACF and PACF plots from stationary data.
3. Conduct an important test of the parameters of the SARIMA model.
4. Diagnostic test model, which is the test of freedom and passivity.
5. Select the best SARIMA models from some of the models already formed based on the smallest AIC values by which the AIC value.
6. Forecasting during the period following intervention using the best SARIMA model obtained at step 5.

After that, calculating residuals, that is, data discrepancies after intervention with predictions and create the residuals graph. Then, perform identification of an intervention response. Next, determines the value of the intervention order of value (b, s, r) by the graph of the intervening response, then conduct a significant test of the parameters of an intervention model, when the judge's parameters are insignificant, we can back to identification other models. If parameters are significant we can diagnostic test of the intervention model, the freedom and autocorrelation test and finally choose the best model based on the smallest AIC for forecasting.

RESULTS AND DISCUSSION

Dercriptive Analysis

A descriptive analysis is done to obtain information about the characteristics of the data used.

Table 6. Descriptive statistics for passenger arrivals

Data	N	Mean	Standard Deviation	Min	Max
The number of arrival passangers	147	127.859	39.724	1.233	196.231

Based on Table 1, the average number of passengers arriving at the Syamsudin Noor period January 2023 is 127.859 with at least 1.233 passengers arriving in one month of May 2020 and the high number of passengers arriving in one month in July 2017. My junction worth 39.724 shows the average count of passengers arriving at Syamsudin Noor airport against the 39.724 arrival rate. Data patterns of the number of passengers arriving at Syamsudin Noor airport from January 2011 to March 2023 can be seen in Figure 2.

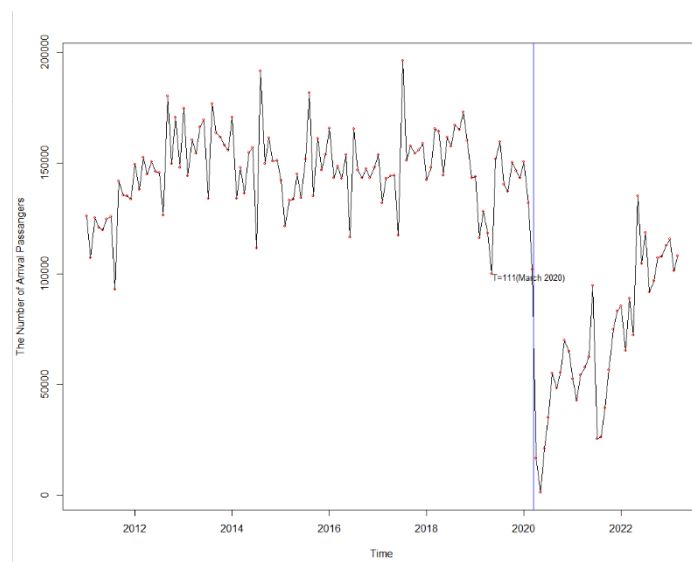


Figure 2. Data pattern the number of arrival passengers with T =111

In figure 2 it is seen that since March 2020 (t =111) the time data pattern changes, with a significant decline in passenger Numbers, and that data is divided into data before the intervention starts from January 2011 - February 2020 (n = 110) and the intervening data from March 2023 (n = 37). The data before an intervention will be used to do the same tools as the intervention.

SARIMA Model before Intervention

1. Data Stationary Test

To find out if the data has been stationary in the raton done by viewing ACF plots and PACF data and ADF tests. ACF plot and PACF data before interventions can be seen on Figure 3.

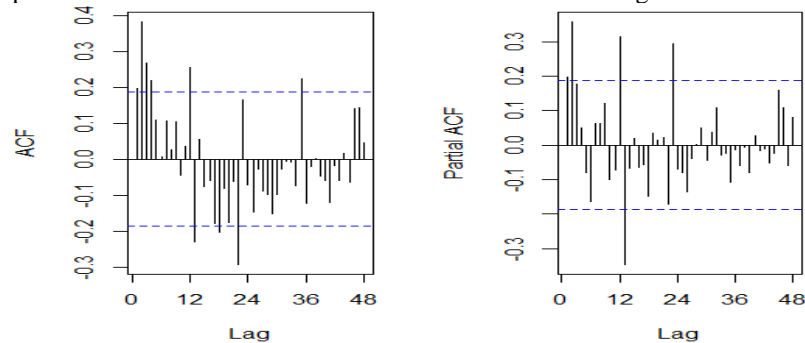


Figure 3. ACF and PACF data patterns before differencing

When viewed from Figure 3, the ACF plot tends to go down slowly so that data can be said to be not stationary in the rate. Aside from looking at ACF's plot, knowing that numerically in the classroom could be done with Augmented Dickey-Fuller (ADF) with the results indicated on Table 2.

Table 1. ADF test data before intervention

	P-value
Augmented Dickey-Fuller test statistic	0,0543

Based on Table 2, b-value test comes from a larger ADF test than the real 5% is 0,0543, which means the data before interventions was not stencumber in ratment, thus needing to be overcome by the process of differencing. After differencing one ($d = 1$), the data before the intervention was rechecked by viewing the ACF plot that can be seen in Figure 4.

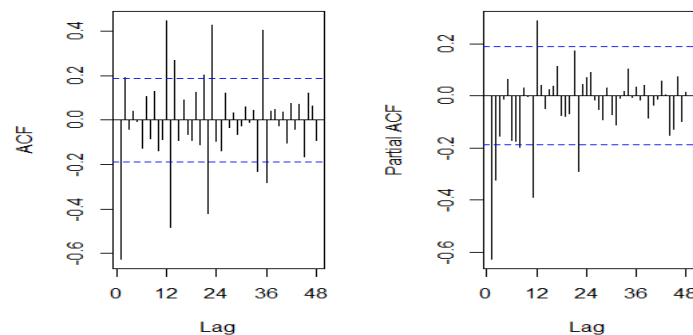


Figure 4. ACF plot and PACF data after 1st differencing

Based on Figure 4 it seems that the ACF plot has a radically downhill pattern, which means data prior to intervention is stationary in ratting. Then the results of the ADF test calculations can be seen at Table 3.

Table 3. ADF test data before intervention

	P-value
Augmented Dickey-Fuller test statistic	0,01

In the ADF test on Table 3, the data before the interventions on differencing 1 ($d = 1$) had a p-value of 0,01 where the value was smaller than a real 5% point which means data can be declared to have been stationary in the rate of Numbers. Further the varying station test was done on the data before the differencing

and got a rounded value on the Box-Cox plot (Figure 5) is 1, it shows that it was stationary in diversity so that no transformation was necessary.

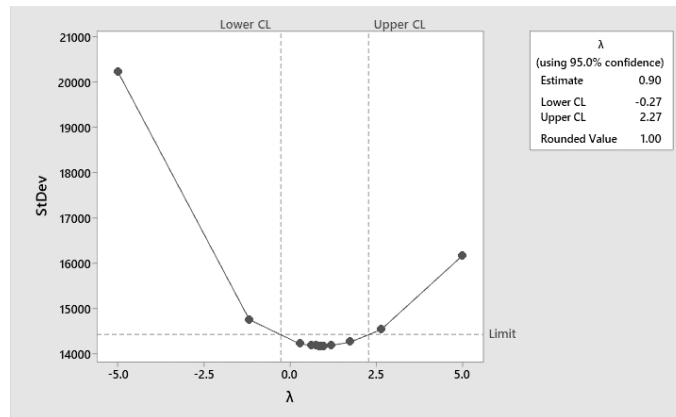


Figure 5. Box-Cox plot before differencing

2. Identification SARIMA Models

Figure 4 shows that the lag on the ACF plot is down significantly to zero after lag 2 and that the plots of PACF are down significantly after lag 2, so that the ARIMA model which is probably AR (2) or MA (2), while the chemical lag on the ACF plot drops slowly after the lag 12 and on the plot of PACF comes down significantly after the lag 12, so that the model is probably the SAR model (1). Based on that, the model tentative is SARIMA(0,1,2)(1,0,0)¹², SARIMA(0,1,1)(1,0,0)¹², SARIMA(2,1,0)(1,0,0)¹² and SARIMA(1,1,0)(1,0,0)¹².

3. Signification Test

Table 4. Test the significance parameters of the sarima model

Model	Coefficient	P-value
SARIMA(0,1,2)(1,0,0) ¹²	MA1($\hat{\theta}_1$)*	-0,824036 <math>< 2.2 \times 10^{-16}</math>
	MA2($\hat{\theta}_2$)	0,163471 0,1049
	SAR1($\hat{\theta}_1$)*	0,464983 1.648×10^{-07}
SARIMA(0,1,1)(1,0,0) ¹²	MA1($\hat{\theta}_1$)*	-0,730824 <math>< 2.2 \times 10^{-16}</math>
	SAR1($\hat{\theta}_1$)*	0,449781 3.851×10^{-07}
SARIMA(2,1,0)(1,0,0) ¹²	AR1($\hat{\phi}_1$)*	-0,796099 <math>< 2.2 \times 10^{-16}</math>
	AR2($\hat{\phi}_2$)*	-0,259379 0,005144
	SAR1($\hat{\theta}_1$)*	0,467986 1.334×10^{-07}
SARIMA(1,1,0)(1,0,0) ¹²	AR1($\hat{\phi}_1$)*	-0,631540 <math>< 2.2 \times 10^{-16}</math>
	SAR1($\hat{\theta}_1$)*	0,500811 4.48×10^{-09}

*) significantly to $\alpha = 0,05$

The value value test results of the four models at Table 4 show that the model of all significant correspondators was the SARIMA(0,1,1)(1,0,0)¹², SARIMA(2,1,0)(1,0,0)¹² and SARIMA(1,1,0)(1,0,0)¹², because the p-value that was gained was smaller than real 5%. The models that all the other significant correspondents next will be tested diagnostic models to see the normality and autocorrelation of data.

4. Diagnostic Test and Choose the Best Model

Table 5. Normality test, autocorrelation and sarima model aic

Model	Jarque-Bera	Ljung-Box	AIC
SARIMA(0,1,1)(1,0,0) ¹²	0,216	0,1308	2413.94
SARIMA(2,1,0)(1,0,0) ¹²	0,0208	0,7138	2412.37
SARIMA(1,1,0)(1,0,0) ¹²	0,009112	0,07969	2417.92

Based on Table 5, then the SARIMA (2,1,0)(1,0,0)¹² is the best model to have the smallest AIC. However, models

SARIMA (2,1,0)(1,0,0)¹² does not satisfy normal assumptions where the value on Jarque-Bera test statistics is $0,0208 < 0,05$. Thus, other models are selected that have the smallest AIC after the SARIMA model (2,1,0)(1,0,0)¹², which is the SARIMA(0,1,1)(1,0,0)¹². The SARIMA model (0,1,1)(1,0,0)¹² has p-value in the Ljung-Box test of 0,1380 which is more than 0,05, so to speak no autocorrelation. For normal testing obtained p-value at the Jarque-Bera test of 0,1308 where the value is more than 0,05 and the assumption of normality is also fulfilled. Then the equation model from SARIMA(0,1,1)(1,0,0)¹² as follows:

$$(1 - 0,449781B^{12})(1 - B)Z_t = (1 + 0,730824B)a_t$$

5. Forecasting as much as Intervention Data

The plot of forecasting results using the SARIMA(0,1,1)(1,0,0)¹² model with actual data can be seen in Figure 6.

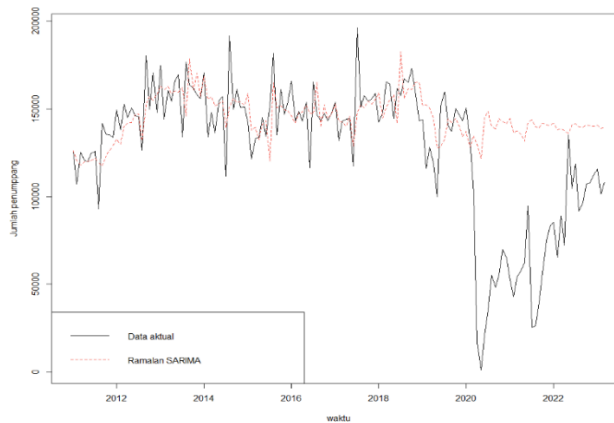


Figure 6. Actual data plot and the sarima model browser

Furthermore, there is an accounting for discrepancies when interventions with forecassions as much as those for interventions using the SARIMA model, which those calculations will be used to create a graph system. The residual graph is used to determine the order on the intervention model.

Intervention Model

1. Identify Intervention Model

The identification of an intervention response is made to determine the order on the intervention model by observing the residual graph. The residual graph can be viewed in Figure 7.

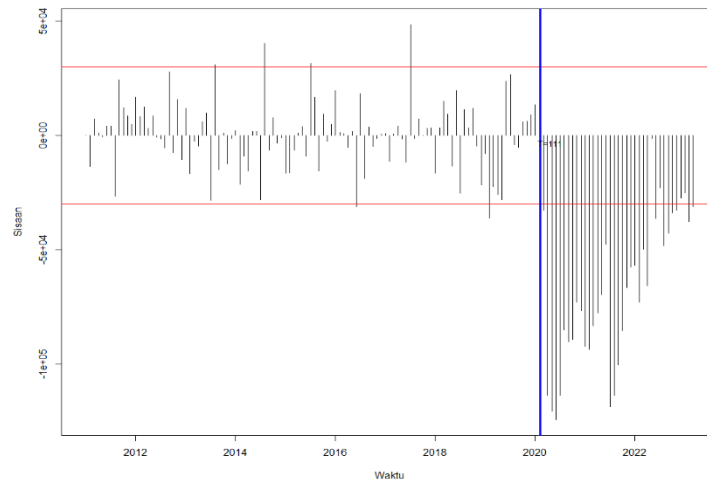


Figure 7. Sarima model residual graph

Figure 7 is seen at the time of the intervention on the 111 observation is a drop that indicates that the response pattern is gradual and permanent, because it is still far beyond the limitation of two standard deviations, so that the function used is the input step.

Residual graph on Figure 7 shows that the number of passengers coming has begun to decline and cut significant threshold in March 2020 ($t = 111$), so selected by the order b is worth 0, as the impact of direct intervention is felt at the time of intervention without any delay. The order of s is believed to be worth 4, and it is seen that for 4 months, from march - June 2020, Indonesia's bilateral growth is still decreasing and is beginning to increase in July 2020, The next order is the order r , the regulatory graph tends to form an exponential pattern so that the order r is worth 1. Therefore, orders from the intervention model based on the sistant graphic identifications $b = 0$, $s = 4$, and $r = 1$. Thus, the tentative model which may be obtained by lowering the order s ($b = 0$, $s = 3$, $r = 1$), ($b = 0$, $s = 2$, and $r = 1$) and ($b = 0$, $s = 1$, and $r = 1$).

2. Signification Test

Table 6. Signification test parameters of intervention models

Model	Coefficient	<i>P</i> -value
$b = 0, s = 4, r = 1$	$\hat{\omega}_0^*$	-56947
	$\hat{\omega}_4^*$	-38482
	$\hat{\delta}_1^*$	0,63792
$b = 0, s = 3, r = 1$	$\hat{\omega}_0^*$	-51581
	$\hat{\omega}_3$	-33602
	$\hat{\delta}_1^*$	0,78287
$b = 0, s = 2, r = 1$	$\hat{\omega}_0^*$	-58917
	$\hat{\omega}_2^*$	-60593
	$\hat{\delta}_1^*$	1,00000
$b = 0, s = 1, r = 1$	$\hat{\omega}_0$	-33643
	$\hat{\omega}_1^*$	84519
	$\hat{\delta}_1$	0,03915

*) significantly to $\alpha = 0,05$

Test of the significance parameters of some of the order interventions on Table 6 suggests that the model of all significant parameters for the measurement $=0,05$ was a model with orders ($b=0, s = 4, r = 1$), and ($b=0, s = 2, r = 1$) because the p -value gained is smaller than real 5%. In the order models ($b = 0, s = 3, r = 1$) and ($b = 0, s = 1$), not all demonstrators were significant at real levels 5% so that models with orders ($b = 0, s = 3, r = 1$) and ($b = 0, s = 1$) could not be used. The two models that all of the other significant parameters will then be seen at its smallest AIC value to be selected as the best model for forecasting.

3. Choosing the Best Model and Diagnostic Test

Table 7. AIC value of intervention models

Model	AIC
$b = 0, s = 4, r = 1$	3234,148
$b = 0, s = 2, r = 1$	3276,424

From Table 7, it can be seen that the model for intervention with orders ($b = 0, s = 4, r = 1$) has a smaller AIC value than the intervening model with the order ($b = 0, s = 2, r = 1$). So that the SARIMA model $(0,1,1)(1,0,0)^{12}(0,4,1)$, is virtually the best model.

At the waste freedom test, it shows that the model SARIMA model $(0,1,1)(1,0,0)^{12}(0,4,1)$, has a p -value Ljung-Box of 0,1761 which has a value of 0,05, so a residual can be said to be free or not self-correlation. Normality test are done with Kolmogorov-Smirnov test, and it's got a p -value of 0,150 where the value is significant on alpha 0,05, so the normal assumption is also fulfilled.

Since residual is normally distributed and there is no autocorrelation, the model of interference with the order $b = 0, s = 4, r = 1$ can be used to over the number of passengers arriving at the airport of Syamsudin Noor. The equation can be formed from model SARIMA $(0,1,1)(1,0,0)^{12}(0,4,1)$, as follows:

$$Y_t = \frac{(-56947+38482B)}{(1-0,63905B)} S_t^{(111)} + \frac{(1+0,730824B)}{(1-0,449781B^{12})(1-B)} a_t$$

$$\text{with } S_t^{(T)} = \begin{cases} 0, & t < 111 \\ 1, & t \geq 111 \end{cases}$$

4. Forecasting

After getting the best model of intervention, there was a forecasting of December 2022-July 2023. Following are the forecasts of the number of passengers arriving at Syamsudin Noor airport of December 2022-July 2023 using an intervention model with order $b = 0, s = 4, r = 1$.

Table 8. The forecasting results with the best intervention model

Periode	Peramalan
April 2023	95.326
Mei 2023	124.462
Juni 2023	105.287
Juli 2023	123.795

From Table 8, the impact of interventions generate predictions with volatile values, but tend to increase. The number of passengers arriving at Syamsudin Noor airport in March 2023 of 108.001 degenerated in April 2023 to 95.326 passengers. However, in May 2023 saw a rise to 124.462 passengers, and in July it became 123.759. Although by June 2023 his values had declined, but the value did not decrease significantly from the actual value in March 2023. To see an image of the increase and decrease in the number of passengers can be seen on the data plot by the forecasting the number of passengers arriving at the Syamsudin Noor airport with an intervention model in December 2022-July 2023 presented with actual data and browser results using SARIMA on the data prior to the intervention (Figure 8).

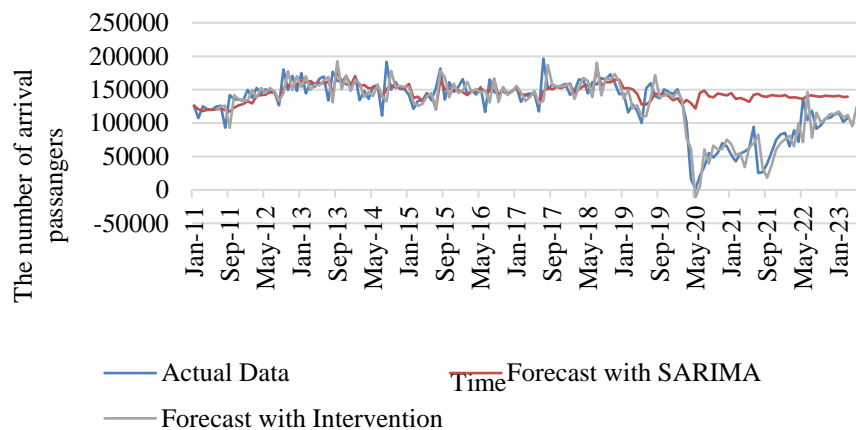


Figure 8. The actual data and predictions using the sarima model and intervention

When viewed from Figure 8, the calculation of the number of passengers coming with the intervention model has a pattern that is said to be increasing and starting again to the point before the Covid-19 pandemic, but the pattern formed does not exceed the previous Covid-19 pandemic. In order to see the ratio between SARIMA's model and the intervention in its forecastal, RMSE could be calculated. The results of both model RMSE calculations can be seen at Table 9.

Table 9. RMSE value of SARIMA and intervention models

Model	RMSE
SARIMA	40.379
Intervensi	20.714

From Table 9, the intervention model has a value of 20.714 less than the SARIMA RMSE, which is 40.379. So, the forecast from an intervention model is closer to actual value than the result from the SARIMA model.

Predictions of incoming passengers at the Syamsudin Noor airport are starting to grow, can affect economic recovery in south kalimantan. Thus, the provincial government of south kalimantan can make efforts that can boost the growth of air transport users, such as developing the potential for tourism and continuing to promote for attracting tourists to south kalimantan.

CONCLUSION

Based on the results of discussions on the browser using the intervention analysis, forecasting for the number of passenger arrivals at Syamsudin Noor Airport Banjarbaru can be modelled with the best model SARIMA(0,1,1)(1,0,0)¹²(0,4,1) with equations: $Y_t = \frac{(-56947+38482B)}{(1-0,63905B)} S_t^{(111)} + \frac{(1+0,730824B)}{(1-0,449781B^{12})(1-B)} a_t$. The modeling of interventions suggests that the impact of interventions is predictive in terms of increasing value.

REFERENCES

- [1] Adisasmita, S.A, Transportasi dan Pengembangan Wilayah, Yogyakarta: Graha Ilmu (2011).
- [2] Sukur, M., Kurniadi, B., Haris & N, F. R. Penanganan Pelayanan Kesehatan di Masa Pandemi Covid-19 dalam Perspektif Hukum Kesehatan. Jurnal Incio Legis, (2020), pp. 1-17.
- [3] Kamaluddin, R, Ekonomi Transportasi: Karakteristik, Teori dan Kebijakan, Jakarta: Ghalia Indonesia (2003).
- [4] Frechtling, D.C, Forecasting Tourism Demand: Methods and Strategies. 1st Edition. London: Routledge (2011).
- [5] Box, G. E., Jenkins, G. M., Reinsel, G. C. & Ljung, G. M, Time Series Analysis; Forecasting and Control, 5th edition, New Jersey (US): John Wiley & Sons (2016).
- [6] Kitagawa, G, Introduction to Time Series Modelling, Japan: Taylor & Francis Group (2010)
- [7] Hanke, J. E. & Winchern, D, Business Forecasting. 9th edition, United States of America: Pearson Education Limited (2014).
- [8] Wei, W. W. S, Time Series Analysis Univariate and Multivariate Methods, 2nd edition, United States of America: Pearson Education, Inc (2006).
- [9] Gujarati, D. N, Basic Econometrics, 4th edition, New York (US): McGraw-Hill (2003).
- [10] Draper, N. & Smith, H, Analisis Regresi Terapan, Jakarta: PT. Gramedia (1992).
- [11] Agresti, A, Statistical Methods for The Statistical Science, 5th edition, United States of America: Pearson Education, Inc (2016).
- [12] Hyndman, R. J. & Athanasopoulos, G, Forecasting: Principles and Practice, 2nd edition, Melbourne: Otexts (2018).
- [13] Mantgomery, D.C., Jennings, C.L. & Kulahci, M, Introduction to Time Series Analysis and Forecasting, 2nd edition, New Jersey: John Wiley & Sons (2015).
- [14] Daniel, W.W, Applied Nonparametric Statistics. 2nd edition, United States: Georgia State University (1998).
- [15] Aritara, R, Analisis Intervensi Fungsi Step Pada Kenaikan Tarif Dasar Listrik (TDL) Terhadap Besarnya Pemakaian Listrik, Yogyakarta: Universitas Negeri Yogyakarta (2011).

Synthetic Minority Oversampling Technique Implementation on Random Forest in Archipelagic Fishing Port of Pemangkat

Joannes Fregis Philosovio Anugrahnu⁽¹⁾, Naomi Nessyana Debatara^{*(1)},
Dadan Kusnandar⁽¹⁾

⁽¹⁾Department of Statistics, Tanjungpura University, Jalan Hadari Nawawi, Pontianak, Indonesia

*Corresponding author: naominessyana@math.untan.ac.id

Abstract. Data Mining provides one of the classification methods called Random Forest. One of the problems in classification is the problem of imbalance class data. Imbalance class data makes the classification results biased towards the majority class. Synthetic Minority Oversampling Technique (SMOTE) can be used to handle the problem. SMOTE produces synthesized data in the minority class with diverse characteristics but still similar to the original data. This study aims to classify the value of measured fishing production at PPN Pemangkat through Random Forest classification using SMOTE to handle the problem of imbalance class data and compare the results with those that do not use SMOTE. This study uses four predictor variables which include: type of fishing gear (X_1), number of days at sea (X_2), number of crew (X_3), and total weight of fish (X_4) with production value as the response variable (Y). The results showed that SMOTE successfully handled the problem of class data imbalance in Random Forest classification. Specificity increased from 0.68% to 5.47%, Precision from 78.98% to 79.43%, and G-Mean from 8.23% to 23.12%. Accuracy and Recall that decreased by 0.73% and 2.2% respectively can be ignored due to the small amount so that Random Forest classification with SMOTE is better than without SMOTE

INTRODUCTION

Random Forest is an algorithm formed from a collection of structured classification trees [1]. Random Forest uses randomization and the Ensemble Learning process so as to reduce the correlation between trees, handle large and diverse amounts of data, and be able to handle missing data problems [2]. One of the problems in data mining classification is the imbalance class data problem. Imbalance class data occurs when the data classes used do not have the same number of data [3]. Classification results on data with imbalance class data problems will be biased towards the majority class, so almost all classification results can certainly always produce the majority class.

A method that can be used in dealing with the problem of imbalance class data is the Synthetic Minority Oversampling Technique (SMOTE). SMOTE can improve classification accuracy by creating synthetic data on minority classes [4]. The synthetic data created by SMOTE basically has a wide variety of data with data characteristics that are close to the original data. This characteristic is an advantage of SMOTE over the usual undersampling and oversampling methods, so that its use can be applied to various types of data. Data on fisheries production is one of the data that does not escape the problem of imbalance class data. In practice, not all fisheries production can have the same results in a certain period of time. Thus, data on fisheries production is suitable for SMOTE implementation.

The Ministry of Maritime Affairs and Fisheries (MMAF) is a government agency that oversees marine and fisheries. There are several types of ports that are under regulation of MMAF, including the Pelabuhan Perikanan Samudera (PPS) and the Pelabuhan Perikanan Nusantara (PPN). There are seven PPS and 18 PPNs spread across Indonesia. Pemangkat PPN in Sambas Regency is one of those PPNs. Pemangkat PPN is one of the centers of fisheries activities in West Kalimantan [5]. Things related to fisheries production such as annual data on fisheries production is one aspect that is considered by Pemangkat PPN. Therefore, this research aims to acquire the classification result of Pemangkat PPN fisheries production using Random Forest with the implementation of SMOTE and compare the result to which the SMOTE is not used in Random Forest.

RESEARCH METHODOLOGIES

Random Forest was developed by Leo Breiman from the Bootstrap Aggregating (Bagging) process. If Bagging randomizes the sample data, Random Forest does the same thing and also randomizes the independent variables used. The results of the decision tree formed then have different sizes and shapes [6]. There are several steps that need to be taken to construct Random Forest classification model [7].

1. Divide the data into training and testing data.
2. In the training data, a number of B sample sets are formed through the bootstrapping process.
3. After forming a number of B sample sets, a Decision Tree is formed on each sample set formed.
4. In each Decision Tree formed, only m variables are used. To select m variables in each Decision Tree, the criterion $1 \leq m \leq \sqrt{p}$ is used where p is the total number of variables.
5. Perform majority voting on all Decision Tree to determine the final classification result (aggregating process).

According to Breimann, the next step is to determine the relative importance of each variable in the Decision Tree from the bootstrapping sample set. The relative importance values are then averaged to see the relative variable importance in Random Forest as a whole. Given each sample set b of B population with the j th variable, error improvement of e_t , and total number of nodes T_b with each of internal node t , relative variable importance level can be calculated from the following equation.

$$I_j^2(b) = \sum_{t=1}^{T_b} e_t^2 I(v(t)_b = j) \quad (1)$$

The overall relative variable importance level then can be calculated through the equation below given the population B for each j th variable in each b tree.

$$I_j^2 = \frac{1}{B} \sum_{b=1}^B I_j^2(b) \quad (2)$$

Classification models in Data Mining, especially machine learning algorithms, uses confusion matrix to evaluate their model's accuracy through performance model indicator. Performance model indicators that are widely used include Accuracy, Recall, Specificity, and Precision. In handling imbalance class data, thus G-mean score is used. To calculate G-mean score, Recall and Specificity are required simultaneously. The following table shows the confusion matrix with its components [8].

Table 1. Confusion Matrix

Class Data	Actual Positive (Majority)	Actual Negative (Minority)
Positive Prediction (Majority)	True Positive (TP)	False Positive (FP)
Negative Prediction (Minority)	False Negative (FN)	True Negative (TN)

Accuracy is a simple performance model indicator by calculating the correct prediction compared to the total sample. Recall is used to measure the accuracy of classification for positive data classes. In the case of imbalance class data, Accuracy and Recall cannot be directly interpreted to explain the model because there is a bias that occurs in the majority class so that the correct prediction in the calculation cannot represent the actual class. Thus, we need to mainly use Precision, Specificity, and G-Mean. Precision and Specificity are used to measure the accuracy of the model in predicting positive classes and negative classes respectively. G-Mean then is used to measure the balance of classification in classifying the majority and minority classes from Precision and Recall. G-Mean score is always positive. The following equations are used to calculate Accuracy, Recall, Specificity, Precision, and G-Mean. All of the components to calculate these equations are generated from confusion matrix.

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \quad (3)$$

$$Recall = \frac{TP}{TP+FN} \quad (4)$$

$$Specificity = \frac{TN}{TN+FP} \quad (5)$$

$$Precision = \frac{TP}{TP+FP} \quad (6)$$

$$G\text{-Mean} = \sqrt{(Recall \times Specificity)} \quad (7)$$

Synthetic Minority Oversampling Technique (SMOTE) is an oversampling technique to increase minority class data to overcome the problem of imbalance class data. The process of increasing data to minority class data is done by generating synthetic data in the training data. Synthetic data is generated using the K-Nearest Neighbor approach. The K-Nearest Neighbor approach determines the distance between the values or scores of the variables used [9]. Value Difference Metric (VDM) is used to determine the distance value between scores of categorical data in K-Nearest Neighbor approach [10]. VDM is a measurement technique to quantify the values of different attributes in a dataset. Given the variables X_p and X_q for each

score of r and C total score of variable for each r independently as well as on their respective variable of p and q , the distance between scores of variables can be calculated using the following equation.

$$\delta(u_p, u_q) = \sum_{r=1}^K \left| \left(\frac{c_{pr}}{c_p} \right) - \left(\frac{c_{qr}}{c_q} \right) \right| \quad (8)$$

Given acquiring the minimum VDM from equation (8) within each j th variable from minority class data of i th data along with β random number from 0 to 1, the synthetic data can be generated through using equation below.

$$m_{ij}^{syn} = m_{ij} + (VDM^{min} - m_{ij})\beta \quad (9)$$

SMOTE is implemented after the data split has conducted mainly aiming on the training set data to be increased using equation (8) and (9). After doing so, the main process of Random Forest classification then continued using the new minority class data generated by SMOTE.

RESULTS AND DISCUSSION

The data is acquired from Pemangkat PPN and is the data of 2021 fisheries production value consist of five variables with four predictor and one response variable. Fishing gear types (X_1), day of trip (X_2), number of crew (X_3), and total weight of fish (X_4) are the predictor variables while production value (Y) is the response variable. Data scoring with each details is shown on Table 2 [11]. The data then is analyzed using Random Forest classification followed by the comparison between before and after Synthetic Minority Oversampling Technique (SMOTE) is implemented to it.

Total data is 3445 and Table 2 shows that there is an imbalance class data problem in the data used. In the production value response variable, it is obtained that the class is divided into two and the production value of more than or equal to 30 million Rupiah is the minority class. It is also obtained that the ratio of majority and minority classes in the data used is 20:80. Only 20% of all measured fishing production values in 2021 were able to reach a number equal to or above IDR 30 million. Thus, the implementation of Synthetic Minority Oversampling Technique (SMOTE) is the right step to take.

The quantity of data in each class of variables used is also uneven. In the predictor variable of fishing gear types, small purse seine is the class that has the largest amount of data compared to other classes in the same variable. This is also the case for the classes on the variables day of trip, number of crew, and total weight of fish. This information reflects that the measured fishing conditions at Pemangkat PPN in 2021 are skewed towards certain characteristics. Therefore, the results obtained from the Random Forest classification method can help the process of increasing the effectiveness of measured fishing at Pemangkat PPN due to its capability to generate diverse model of tree simultaneously.

Table 2. Scoring and Data Descriptive

Variables	Score	Description	Quantity
Production Value	1	< IDR 30 Million	2730
	2	≥ IDR 30 Million	715
Fishing Gear Types	1	Small purse seine	3337
	2	Stick held dip net	16
	3	Cast net	64
	4	Bottom long line	28
Day of Trip	1	< 21 days	713
	2	21 to 40 days	2648
	3	> 40 days	84
Number of Crew	1	< 11 people	44
	2	11 to 20 people	1369
	3	> 20 people	2032
Total Weight of Fish	1	< 10 Tons	336
	2	≥ 10 Tons	3109

In this research, the data split is on the ratio on 80:20 with 80% for training data and 20% for testing data. The data split is followed by the SMOTE implementation on increasing the minority class data as the data is already confirmed as having imbalance class data problem. Figure 1 is showing the initial split data side by side with the final split data after SMOTE already implemented.

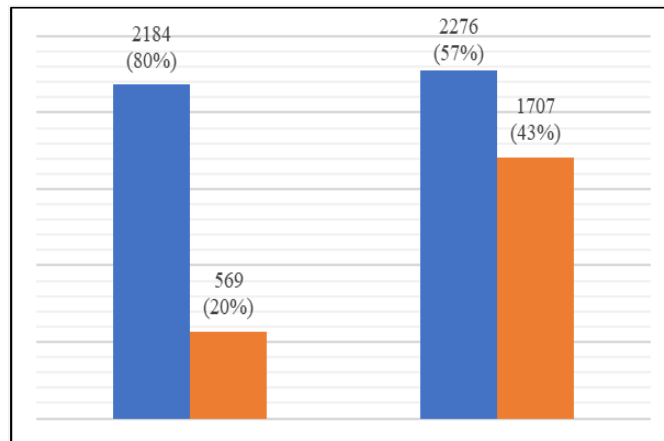


Figure 1. Split Data Before and After SMOTE Implementation

The next step is to use Random Forest to do classification on the both split data. Number of trees created is 500 with all variables available in each tree's split node. After acquiring the model, the result of average relative variable importance is also generated. Figure 2 is showing the side by side comparison of average relative variable importance for both Random Forest model with and without implementation of SMOTE. The results are similar thus proving that SMOTE implementation does not affect the characteristics of the original dataset.

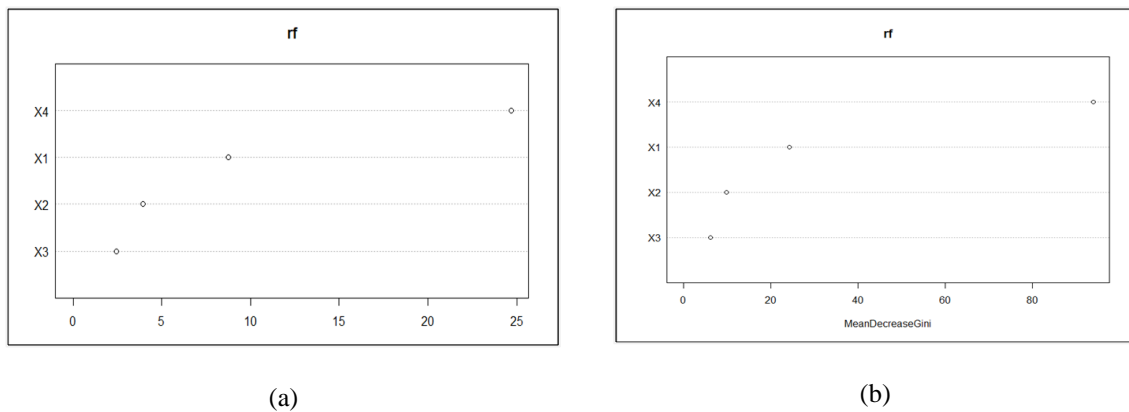


Figure 2. Average Relative Variable Importance Comparison

Each score in performance model indicator of Accuracy, Recall, Specificity, Precision, and G-Mean have been obtained for Random Forest classification with and without SMOTE from confusion matrix. The direct comparison of the performance model indicator is shown on Table 3.

Table 3. Performance Model Indicator Comparison

Performance Model Indicator	Random Forest without SMOTE	Random Forest with SMOTE
Accuracy	78.9%	78.17%
Recall	99.81%	97.61%
Specificity	0.68%	5.47%
Precision	78.98%	79.43%
G-Mean	8.23%	23.12%

Random Forest classification with SMOTE provides better results than Random Forest classification without SMOTE. The main difference lies in the Specificity, Precision, and G-Mean scores which have increased by 4.79%, 0.45%, and 14.89% respectively. Random Forest classification with SMOTE has better classification accuracy for minority classes. This difference proves that the bias arising in the problem of imbalance class data towards the majority data class can be mitigated by the implementation of SMOTE. In

addition, the decreased Accuracy and Recall values are negligible in the case of imbalance class data considering that the bias in the majority class greatly affects the scores generated by Accuracy and Recall.

CONCLUSION

Synthetic Minority Oversampling Technique (SMOTE) can be implemented well in the classification of measured fishing production data from the Pemangkat Pelabuhan Perikanan Nusantara (PPN) in 2021 with the Random Forest classification. SMOTE adds a layer of analysis that is useful in identifying and being a solution to the problem of imbalance class data in Random Forest classification conducted in the research. The results of Random Forest classification with the implementation of SMOTE also significantly helping on handling imbalance class data problem. There is an increase in Specificity performance indicators by 4.79%, Precision by 0.45%, and G-Mean by 14.89% Therefore, SMOTE is proven to be able to improve the performance of Random Forest classification in dealing with data class imbalance problems. Accuracy and Recall values were decreased but were negligible due to the bias that occurred in the majority class.

REFERENCES

- [1] Breiman, L., "Random Forests", *Machine Learning*, 45(1), 5–32 (2001).
- [2] Ali, J., Khan, R., Ahmad, N., & Maqsood, I., "Random Forests and Decision Trees", *IJCSI International Journal of Computer Science Issues*, 9(5) (2012).
- [3] Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P., "SMOTE: Synthetic Minority Oversampling Technique", *Journal of Artificial Intelligence Research*, 16, 321–357 (2002).
- [4] Permatasari, R. D., Rizki, S. W., & Debataraja, N. N., "Penerapan Synthetic Minority Oversampling Technique Dalam Mengatasi Data Tidak Seimbang pada Metode Classification and Regression Tree", *Bimaster: Buletin Ilmiah Matematika, Statistika dan Terapannya*, 9(1), 231–238 (2020).
- [5] Safitri, I., & Magdalena, W., "Perikanan Tangkap Purse Seine di Pelabuhan Perikanan Nusantara (PPN) Pemangkat Kalimantan Barat", *Jurnal Laut Khatulistiwa*, 1(3), 89–96 (2018).
- [6] Jatmiko, Y. A., Padmadisastra, S., & Chadidjah, A., "Analisis Perbandingan Kinerja Cart Konvensional, Bagging Dan Random Forest Pada Klasifikasi Objek: Hasil Dari Dua Simulasi", *Media Statistika*, 12(1), 1-12 (2019).
- [7] Lee, T.-H., Ullah, A., & Wang, R., "Bootstrap Aggregating and Random Forest. In P. Fuleky (Ed.)", *Macroeconomic Forecasting in the Era of Big Data*, 52, 389–429 (2020).
- [8] Syukron, M., Santoso, R., & Widiharih, T., "Perbandingan Metode Smote Random Forest dan Smote Xgboost untuk Klasifikasi Tingkat Penyakit Hepatitis C pada Imbalance Class Data", *Jurnal Gaussian*, 9(3), 227–236, (2020).
- [9] Sofyan, S., & Prasetyo, A., "Penerapan Synthetic Minority Oversampling Technique (SMOTE) Terhadap Data Tidak Seimbang pada Tingkat Pendapatan Pekerja Informal Di Provinsi D.I. Yogyakarta Tahun 2019", *Seminar Nasional Official Statistics*, 2021(1), 868–877, (2021).
- [10] Hastie, T., Tibshirani, R., & Friedman, J., *The Elements of Statistical Learning*, New York: Springer, 2009.
- [11] Irnawati, R., Simbolon, D., Wiryawan, B., Murdianto, B., & Nurani, T. W., "Leading commodity analysis of capture fisheries in Karimunjawa National Park", *Jurnal Perikanan dan Kelautan*, 1(1), 11-17 (2011).

Forecasting Reference Coal Price Using Transfer Function Multiple Inputs

Sri Wahyuningsih^{*⁽¹⁾}, Meiliyani Siringoringo⁽¹⁾, Ika Purnamasari⁽¹⁾, Darnah A. Nohe⁽¹⁾,
Ruli Yuniarto⁽²⁾, Jumarni⁽²⁾, and Siti Nur Rahmah⁽²⁾

⁽¹⁾*Statistics Study Program, Mathematics Department, Faculty of Mathematics and Natural Sciences, Mulawarman University, Samarinda, Indonesia*

⁽²⁾*Laboratory of Economy and Business, Mathematics Department, Faculty of Mathematics and Natural Sciences, Mulawarman University, Samarinda, Indonesia*

**Corresponding author: swahyuningsih@fmipa.unmul.ac.id*

Abstract. The transfer function is a forecasting model that pays attention to the dependence of one variable by involving the dependence. This transfer function model is capable of capturing linear data patterns. The purposes of this study are to obtain a transfer function model on the reference coal price in Indonesia and obtain the results of reference coal price forecasting in Indonesia using a transfer function model. The results of the study obtained that the reference coal price in Indonesia is influenced by coal exports with a forecasting accuracy rate MAPE 6.4069% and MAPE 6.2624%. Forecasting with the transfer function model shows that reference coal price in Indonesia for the period May 2023 to April 2024 tends to be stable.

Keywords: forecasting, reference coal price, transfer function, multiple inputs.

INTRODUCTION

The transfer function is a model that describes the predicted value in the future on time series data or is called an output sequence based on the past values of the series and is based on one or more time series related to the output series or is called an input sequence (Hyndman & Athanasopoulos, 2018). The transfer function model is a forecasting model that can be used in periodic series data with more than two variables. The transfer function is divided into two, namely the single input transfer function and the multiple input transfer function (Montgomery, 2008). The single input transfer function describes the predicted value of a time series that has a relationship to the input variable. Meanwhile, in the multi-input transfer function there are several input variables that are included in a model. The main purpose of modelling the transfer function is to determine the role of the indicator (leading indicator) or input sequence in order to determine the output sequence variable (Hanke & Wichern, 2005).

Indonesia is one of the largest coal producing countries in the world. Reference coal price is the price of coal which is the basis for calculating the minimum to pay royalties and state taxes. Price increases accompanied by high demand for coal exports to certain countries can be used as an opportunity for Indonesia to gain large profits for the country. For this reason, reference coal price forecasting is very important for the government and coal entrepreneurs. Several previous studies have also discussed reference coal prices, namely research conducted by Umiyati, et al (2020), regarding reference coal price forecasting using the Support Vector Regression (SVR) method which is optimized with Particle Swarm Optimization (PSO) and Improved-Particle Swarm Optimization (IPSO)). Research that discusses reference coal price and other influencing variables was conducted by Anindita & Syaputra (2017), regarding the analysis of the rupiah exchange rate, coal production, domestic coal demand, and reference coal price on Indonesian coal export volume using the multiple linear regression method. Based on this research, the reference coal price has a linear relationship with coal production, domestic coal demand, coal export volume and the rupiah exchange rate.

The objectives in this study are:

1. Obtain a transfer function model that explains the relationship between coal production, domestic coal sales/demand, coal exports and the IDR to USD exchange rate to RCP in Indonesia.
2. Obtain the RCP forecasting results in Indonesia using the transfer function model.

MATERIALS AND METHODS

Autoregressive Integrated Moving Average (ARIMA) Model

According to Aswi and Sukarna (2006), the ARIMA model was introduced in 1970 by Box and Jenkins through their book entitled Time Series Analysis. ARIMA is often also called the Box-Jenkins time series method. ARIMA has very good accuracy for both short-term and long-term forecasting. ARIMA can be interpreted as a combination of two models, namely the AR model which is integrated with the MA model. In general, the ARIMA (p,d,q) model can be written as follows:

$$\phi_p(B)(1-B)^d Z_t = \theta_q(B)e_t \quad (1)$$

where:

B : backshift operator

d : differencing ordo

$$\begin{aligned} \phi_p(B) &= 1 - \phi_1(B) - \phi_2(B^2) - \dots - \phi_p(B)^p \\ \theta_q(B) &= 1 - \theta_1(B) - \theta_2(B^2) - \dots - \theta_q(B)^q \end{aligned}$$

(Wei, 2006).

Transfer Function

In general, the output sequence may be affected by several input sequences, so the causal model for the multi-input transfer function is:

$$Y_t = \sum_j^k v_j(B)X_{jt}^* + n_t \quad (2)$$

$$Y_t = \sum_j^k \frac{\omega_j(B)}{\delta_j(B)} B^{bj} X_{jt}^* + \frac{\theta(B)}{\phi(B)} a_t^* \quad (3)$$

where $v_j(B) = \frac{\omega_j(B)B^{bj}}{\delta_j(B)}$ is the transfer function for the j th input series X_{jt}^* and a_t^* is assumed to be independent for each input series X_{jt}^* , $j = 1, 2, \dots, k$. Transfer function response weights $\frac{\omega_j(B)B^{bj}}{\delta_j(B)}$ for each input variable are defined in the single input transfer function model (Wei, 2006).

Pre-Whitening of Input and Output Series

Pre-whitening of the input series is done to simplify the input series by eliminating all known patterns so that only white residuals remain. The pre-whitening of the X_t^* input series with the ARIMA($p_x, 0, q_x$) process is:

$$\phi_x(B)X_t^* = \theta_x(B)\alpha_t \quad (4)$$

Change the X_t^* α input series to α_t series as follows:

$$\alpha_t = \frac{\phi_x(B)}{\theta_x(B)} X_t^* \quad (5)$$

If pre-whitening is done for X_t^* then pre-whitening is also applied to Y_t^* in order to maintain the integration of functional relationships. So generated

$$\begin{array}{ccc} \text{Input } X_t^* & \xrightarrow{\text{transfer function}} & \text{output } Y_t^* \\ \text{Input } X_t^* \frac{\phi_x(B)}{\theta_x(B)} & \xrightarrow{\text{transfer function}} & \text{output } Y_t^* \frac{\phi_x(B)}{\theta_x(B)} \end{array}$$

The transformation on Y_t^* doesn't have to change X_t^* to a white residual. The following is the pre-whitening time series Y_t^* which has been pre-whitened to be called β_t :

$$\beta_t = \frac{\phi_x(B)}{\theta_x(B)} Y_t^* \quad (6)$$

(Makridakis, et al., 1999)

RESULTS AND DISCUSSION

The data collected is reference coal price data, realized coal production, realized domestic coal sales/demand, coal exports and the IDR-USD exchange rate in Indonesia through the official website of the Ministry of Energy and Mineral Resources, https://www.minerba.esdm.go.id/harga_acuan, the official

Minerba Satu Data Indonesia website, <https://modi.esdm.go.id/production-batubara> and the Indonesian Ministry of Trade's Satu Data official website, <https://satudata.kemendag.go.id/data-information/domestic-trade/exchange-rates>. Figures 1 to 7 are time series plots of these data. These are monthly data from January 2016 to April 2023.

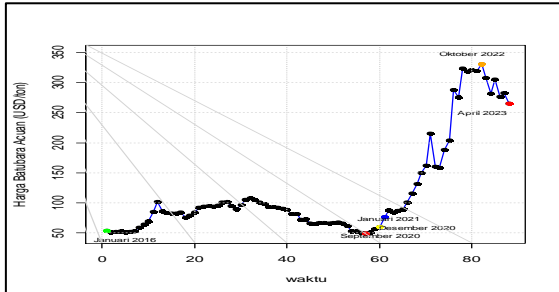


Figure 1. Reference coal price

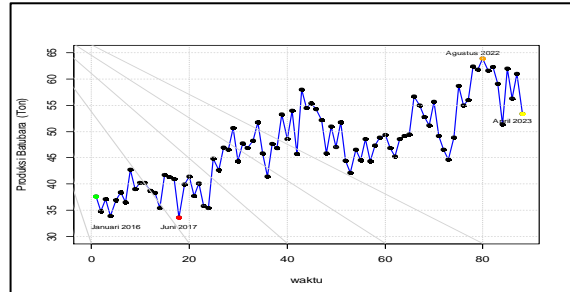


Figure 2. Realized coal production

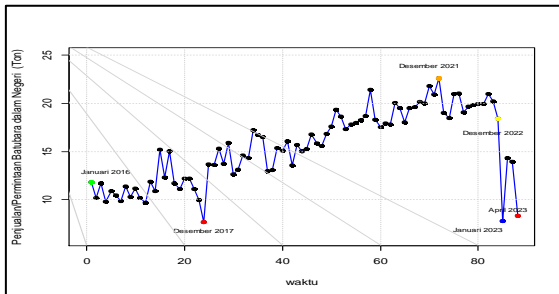


Figure 3. Realized coal sales/demand

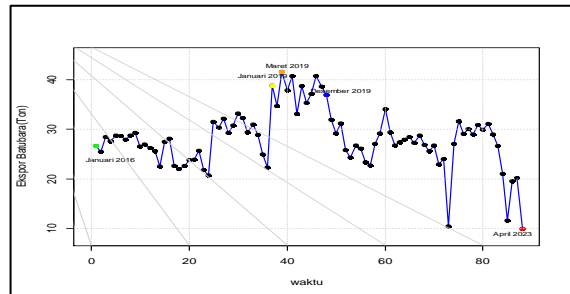


Figure 4. Coal export

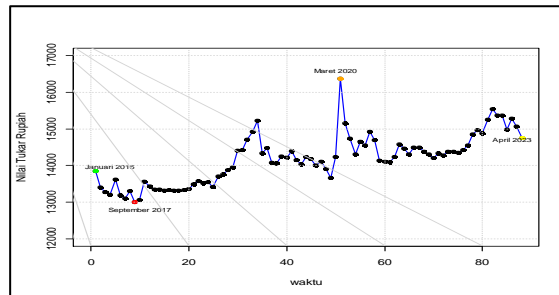


Figure 5. IDR- USD exchange rate

Test of Independence Between Variables

In multivariate time series analysis, it is necessary to test the dependency (correlation analysis) between variables, namely the test is carried out to see the effect between the reference coal price, coal production, domestic coal sales/demand, coal exports and the IDR-USD exchange rate. Dependency testing can be done by using the Bartlett of Sphericity test to prove whether the variables are related to each other or not. The Bartlett of Sphericity testing hypothesis is as follows:

$$R = \begin{bmatrix} r_{YY} & r_{YX_1} & r_{YX_2} & r_{YX_3} & r_{YX_4} \\ r_{X_1Y} & r_{X_1X_1} & r_{X_1X_2} & r_{X_1X_3} & r_{X_1X_4} \\ r_{X_2Y} & r_{X_2X_1} & r_{X_2X_2} & r_{X_2X_3} & r_{X_2X_4} \\ r_{X_3Y} & r_{X_3X_1} & r_{X_3X_2} & r_{X_3X_3} & r_{X_3X_4} \\ r_{X_4Y} & r_{X_4X_1} & r_{X_4X_2} & r_{X_4X_3} & r_{X_4X_4} \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0,6888 & 0,3292 & -0,2725 & 0,5500 \\ 0,6888 & 1 & 0,6150 & 0,2129 & 0,7203 \\ 0,3292 & 0,6150 & 1 & 0,1607 & 0,6220 \\ -0,2725 & 0,2129 & 0,1607 & 1 & -0,0355 \\ 0,5500 & 0,7203 & 0,6220 & -0,0355 & 1 \end{bmatrix}$$

$$\begin{aligned}\chi^2 &= -\left(n - 1 - \frac{2m - 5}{6}\right) \ln|\mathbf{R}| \\ &= -\left(83 - 1 - \frac{2(5) - 5}{6}\right) \ln|0,0744| \\ &= 223,8653\end{aligned}$$

The calculation results show that the value $\chi^2 = 223,8653$ is greater than $\chi_{\alpha,df}^2 = 18,037$ with $\alpha = 0,05$ and $df = \frac{1}{2}m(m - 1) = 10$ where $m = 5$ which means H_0 is rejected. It can be concluded that there is a correlation between reference coal price, coal production, domestic coal sales/demand, coal exports and the IDR-USD exchange rate.

ARIMA Model of Input and Output Series

Identification of the temporary ARIMA model is carried out by looking at the autocorrelation function and partial autocorrelation function graphs. The autocorrelation function value for reference coal price is cut off after lag 2 and the partial autocorrelation function value is cut off after lag 1, in order to obtain a temporary ARIMA model which can be seen in Table 1.

Table 1. ARIMA Model of Reference Coal Price

Model	
ARIMA (0,1,1)	$Y_t^* = Y_{t-1}^* + e_t - \theta_1 e_{t-1}$
ARIMA (1,1,0)	$Y_t^* = (1 + \phi_1)Y_{t-1}^* - \phi_1 Y_{t-2}^* + e_t$
ARIMA (1,1,1)	$Y_t^* = (1 + \phi_1)Y_{t-1}^* - \phi_1 Y_{t-2}^* + e_t - \theta_1 e_{t-1}$
ARIMA (0,1,2)	$Y_t^* = Y_{t-1}^* + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2}$
ARIMA (1,1,2)	$Y_t^* = (1 + \phi_1)Y_{t-1}^* - \phi_1 Y_{t-2}^* + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2}$

A significant ARIMA model is an ARIMA model whose parameters have a p-value less than 0.05. Based on the parameter significance test in Table 2, there are several models whose parameters have a p-value smaller than 0.05, namely the ARIMA (0,1,1) and ARIMA (1,1,0) models, so it was decided that H_0 was rejected. It was concluded that the significant models for reference coal price were ARIMA (0,1,1) and ARIMA (1,1,0).

Table 2. Parameter Estimation of ARIMA Model for Reference Coal Price

Model	Parameter	Estimasi	t	df	$t_{0,025;df}$	p-value	Decision
ARIMA (0,1,1)	θ_1	0,3252	3,5342	87	1,9876	0,0006	H_0 is rejected
ARIMA (1,1,0)	ϕ_1	0,3976	4,0629	87	1,9876	0,0001	H_0 is rejected
ARIMA (1,1,1)	θ_1	0,5941	3,0831	86	1,9879	0,0027	H_0 is rejected
	ϕ_1	-0,2358	-1,0172	86	1,9879	0,3119	H_0 is not rejected
ARIMA (0,1,2)	θ_1	0,3320	3,2114	86	1,9879	0,0019	H_0 is rejected
	θ_2	0,1996	1,7785	86	1,9879	0,0788	H_0 is not rejected
ARIMA (1,1,2)	ϕ_1	0,5436	1,6729	85	1,9883	0,0979	H_0 is not rejected
	θ_1	-0,1899	-0,5653	85	1,9883	0,5733	H_0 is not rejected
	θ_2	0,0362	0,2134	85	1,9883	0,8315	H_0 is not rejected

The model meets the assumptions of normal distribution of residual testing if the p-value is > 0.05 . Based on the residual normality test in Table 3, it was found that the p-value of the ARIMA (0,1,1) and ARIMA (1,1,0) models was greater than 0.05 so it was decided that H_0 failed to be rejected. It was concluded that the residual data of the ARIMA (0,1,1) and ARIMA (1,1,0) models came from populations that were normally distributed.

Table 3. Test of Normality for Residual for Reference Coal Price Model

Model	D	$D_{0,025:88}$	p-value	Decision
ARIMA (0,1,1)	0,0984	0,1480	0,3397	H_0 is not rejected
ARIMA (1,1,0)	0,1174	0,1480	0,1633	H_0 is not rejected

Figure 6.a and 6.b show that the residual ARIMA (0,1,1) and ARIMA (1,1,0) have p-values for all lags above the threshold line or significance level $\alpha = 0.05$, so it is decided that H_0 fails to be rejected. It was concluded that there was no autocorrelation in the residual ARIMA (0,1,1) and ARIMA (1,1,0) models for reference coal price.

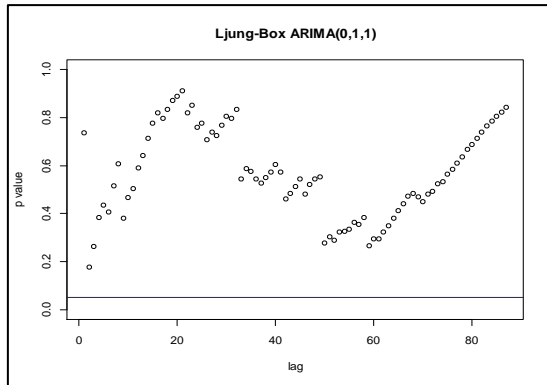


Figure 6.a. Ljung Box test of ARIMA (0,1,1) for reference coal price

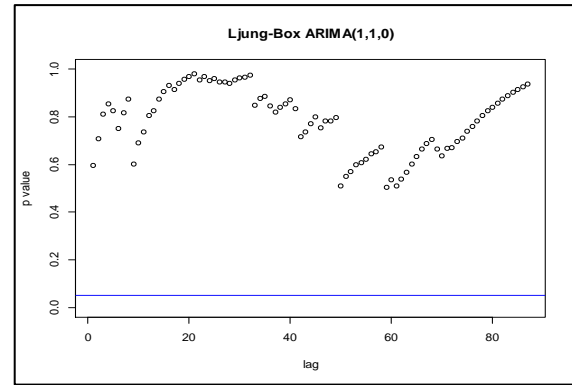


Figure 6.b. Ljung Box test of ARIMA (1,1,0) for reference coal price

Selection of the best model based on the smallest AIC value. The table for selecting the best ARIMA model using AIC criteria can be seen in Table 4. This table shows that the ARIMA(1,1,0) model has the smallest AIC value. The ARIMA(1,1,0) can be stated as follows:

$$Y_t^* = (1 + \phi_1)Y_{t-1}^* - \phi_1 Y_{t-2}^* + e_t$$

$$Y_t = Y_{t-1}^* + 0,3976Y_{t-1}^* - 0,3976Y_{t-2}^* + e_t \quad (7)$$

Table 4. AIC for Reference Coal Price Model

Model	AIC
ARIMA(0,1,1)	-916,06
ARIMA(1,1,0)	-919,54

With the same steps, an ARIMA model is also made for each input variable for coal production (Table 5), realized coal sales/demand (Table 7), coal export (Table 8).

Tabel 6. Parameter Estimation of ARIMA Model for Coal Production

Model	Parameter	Estimasi	t	df	$t_{0,025:df}$	p-value	Decision
ARIMA (0,1,1)	θ_1	-0,5277	-5,5881	87	1,9876	$2,59 \times 10^{-7}$	H_0 is rejected
ARIMA (1,1,0)	ϕ_1	-0,4813	-5,0780	87	1,9876	$2,148 \times 10^{-6}$	H_0 is rejected
ARIMA (1,1,1)	θ_1	-0,1929	-0,7021	86	1,9879	0,4845	H_0 is not rejected
	ϕ_1	-0,3601	-1,2824	86	1,9879	0,2031	H_0 is not rejected

Tabel 7. Parameter Estimation of ARIMA Model for Coal Sales/Demand

Model	Parameter	Estimasi	t	df	$t_{0,025;df}$	p -value	Decision
ARIMA (0,1,1)	θ_1	-0,4588	-4,4733	87	1,9876	$2,317 \times 10^{-5}$	H_0 is rejected
ARIMA (1,1,0)	ϕ_1	-0,3632	-3,5147	87	1,9876	0,0007	H_0 is rejected
ARIMA (1,1,1)	θ_1	0,0491	0,2095	86	1,9879	0,8346	H_0 is not rejected
	ϕ_1	-0,4978	-2,4091	86	1,9879	0,0182	H_0 is not rejected

Tabel 8. Parameter Estimation of ARIMA Model for Coal Export

Model	Parameter	Estimasi	t	df	$t_{0,025;df}$	p -value	Decision
ARIMA (0,1,1)	θ_1	-0,3483	-3,2353	87	1,9876	0,0017	H_0 is rejected
ARIMA (1,1,0)	ϕ_1	-0,3053	-2,9438	87	1,9876	0,0042	H_0 is rejected
ARIMA (1,1,1)	θ_1	0,1198	0,3035	86	1,9879	0,7622	H_0 is not rejected
	ϕ_1	-0,4584	-1,2535	86	1,9879	0,2134	H_0 is not rejected

Based on Figure 7.a it can be seen that there are no cut off lags on the autocorrelation graph and based on Figure 7.b it can be seen that there are no cut off lags on the partial autocorrelation graph, so there is no model that can be used in modeling the ARIMA and transfer function.

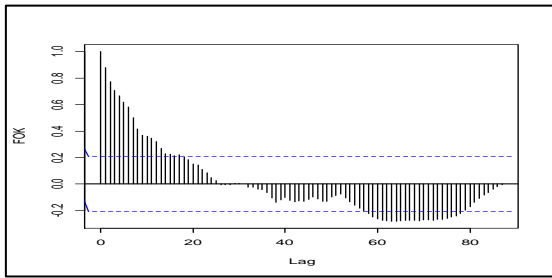


Figure 7.a. Autocorrelation function for IDR-USD exchange rate

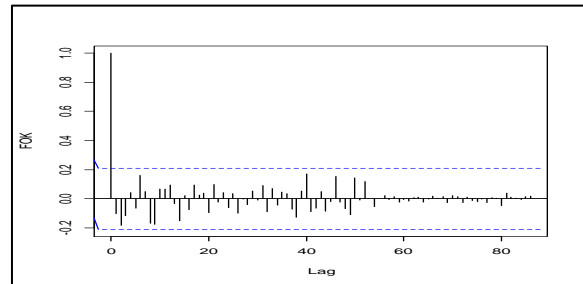


Figure 7.b Partial autocorrelation function for IDR-USD exchange rate

Transfer Function Constuction

Pre-whitening process for input series using the ARIMA model and noise of the model:

$$a_{1t}^* = X_{1t}^* - X_{1t-1}^* - 0,5277a_{1t-1}^* \quad (8)$$

$$a_{2t}^* = X_{2t}^* - X_{2t-1}^* - 0,4588a_{2t-1}^* \quad (9)$$

$$a_{3t}^* = X_{3t}^* - X_{3t-1}^* - 0,3483a_{3t-1}^* \quad (10)$$

Pre-whitening process for output series using the ARIMA model and noise of the model:

$$\beta_{1t} = Y_t - Y_{t-1} - 0,5277\beta_{1t-1} \quad (11)$$

$$\beta_{2t} = Y_t - Y_{t-1} - 0,4588\beta_{2t-1} \quad (12)$$

$$\beta_{3t} = Y_t - Y_{t-1} - 0,3483\beta_{3t-1} \quad (13)$$

Based on Figure 8.a and 8.b, it can be concluded that the order (b,r,s) is (0, 0, 0). Based on Figure 8.c it can be seen that the first significant lag is the first lag, so the order $b = 1$ is obtained. Order $b = 1$ indicates a delay of one period before the coal export input sequence affects the reference coal price (output

sequence). Based on Figure 8.c there is no significant lag after the second lag, then the order $s = 0$ is obtained. Based on Figure 8.c the cross correlation function graph does not form a sine or exponential pattern, then $r = 0$ is obtained. Based on the description above it can be concluded that the order (b,r,s) is $(1, 0, 0)$. The initial model of the transfer function based on Equation (2) is as follows:

$$Y_t = \omega_{01}(X_{1t}) + \omega_{02}(X_{2t}) + \omega_{03}B(X_{3t}) + \eta_t \quad (14)$$

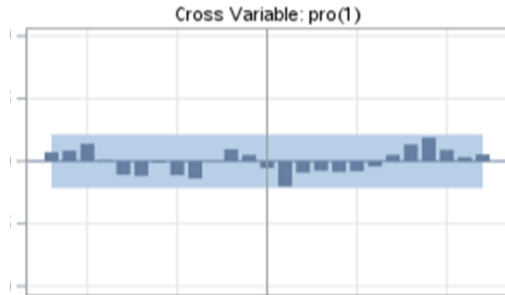


Figure 8.a. Cross Correlation Function Reference Coal Price and Coal Production

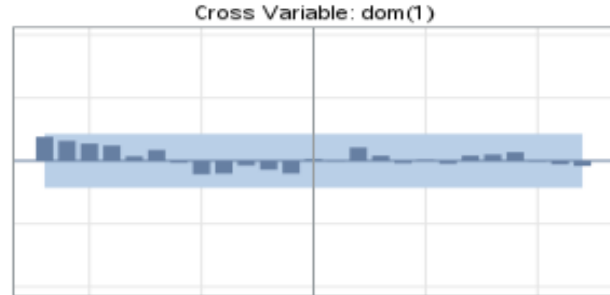


Figure 8.b. Cross Correlation Function Reference Coal Price and Coal Sales/Demand

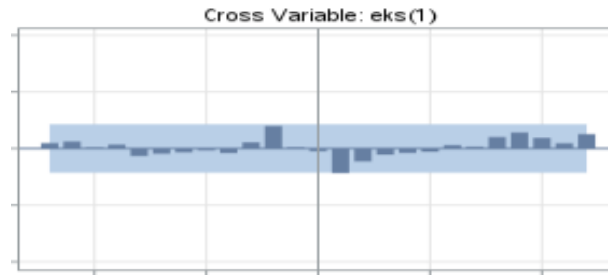


Figure 8.c. Cross Correlation Function Reference Coal Price and Coal Export

Based on Table 9, it can be seen that the initial model of the transfer function of coal exports to reference coal price has a value of $|t| > t_{0,025;87}$ or $p\text{-value} < \alpha$. Meanwhile, coal production and sales/demand for domestic coal have a value of $|t| < t_{0,025;87}$ or $p\text{-value} > \alpha$. It can be decided that the coal export variable H is rejected and it is concluded that the parameter ω_{03} is significant.

Table 9. Parameter Estimation of The Initial Transfer Function

Parameter	Estimation	t	$t_{0,025;87}$	p - value	Decision
ω_{01}	-0,00038	-0,2200	0,9876	0,8274	H_0 is not rejected
ω_{02}	0,00003	0,7800	0,9876	0,4397	H_0 is not rejected
ω_{03}	-0,00001	-2,1900	0,9876	0,0316	H_0 is rejected

The final identification of the transfer function model with two input sequences is carried out by combining the initial model and its residual model based on the ARIMA (0,0,1) and ARIMA (1,0,0) residual series models. The combination of the initial transfer function model with the residual series ARIMA model can be seen in Table 10.

Table 10. The Final Transfer Function

Model	Transfer Function
ARIMA (0,0,1)	$Y_t = \omega_{03}BX_{3t-1} + (1 - \theta_1B)a_t^*$
ARIMA (1,0,0)	$Y_t = \omega_{03}BX_{3t-1} + \frac{a_t^*}{1 - \phi_1B}$

Based on Table 11, it can be decided that H_0 is rejected and it can be concluded that the final model parameters of the ARIMA (0,0,1) and ARIMA (1,0,0) transfer functions are significant. After testing the significance of the parameters, then a diagnostic of the transfer function model is carried out which consists of cross-correlation test and autocorrelation test.

Table 11. Parameter Estimation of The Last Transfer Function

Transfer Function	Parameter	Estimation	t	f ^d	t _{0,025:df}	p-value	Decision
ARI MA (0,0,1)	θ_1	-0,3931	-3,5200	8 6	1,98 79	0,0007	H_0 is rejected
	ω_{03}	-0,00001	-2,5200	8 6	1,98 79	0,0136	H_0 is rejected
ARI MA (1,0,0)	ϕ_1	0,4198	4,2400	8 6	1,98 79	<0,000 1	H_0 is rejected
	ω_{03}	-0,00001	-2,4900	8 6	1,98 79	0,0149	H_0 is rejected

Table 12 shows that the MAPE value of the two models is less than 10% and does not differ much in value. The two transfer function models produce very good forecasts.

Table 12 Accuracy of The Final Transfer Function Model

Transfer Function	MAPE
ARIMA (0,0,1)	6,4069%
ARIMA (1,0,0)	6,2624%

Based on Figure 9.a and 9.b show that the data pattern on the predicted results is close to the actual data pattern and reference coal price forecasting in Indonesia tends to be stable.

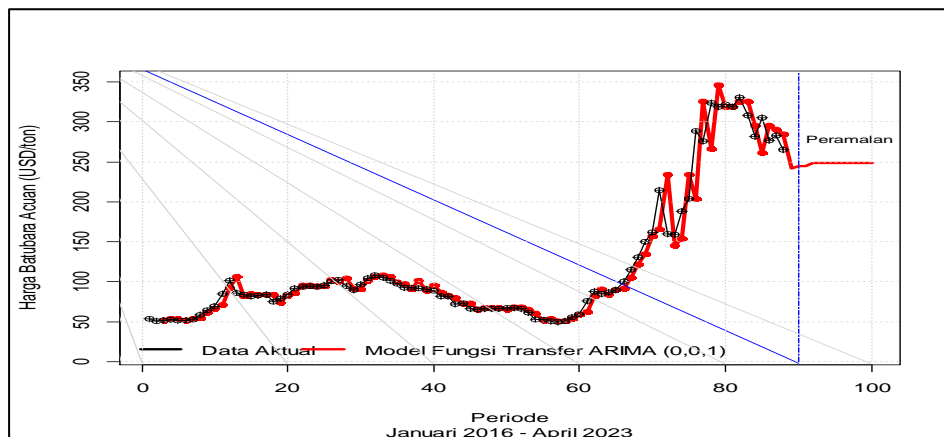


Figure 9.a. Transfer Function ARIMA (0,0,1)

$$\hat{Y}_t = -0,3931Y_{t-1} - 0,00001X_{3t-1}$$

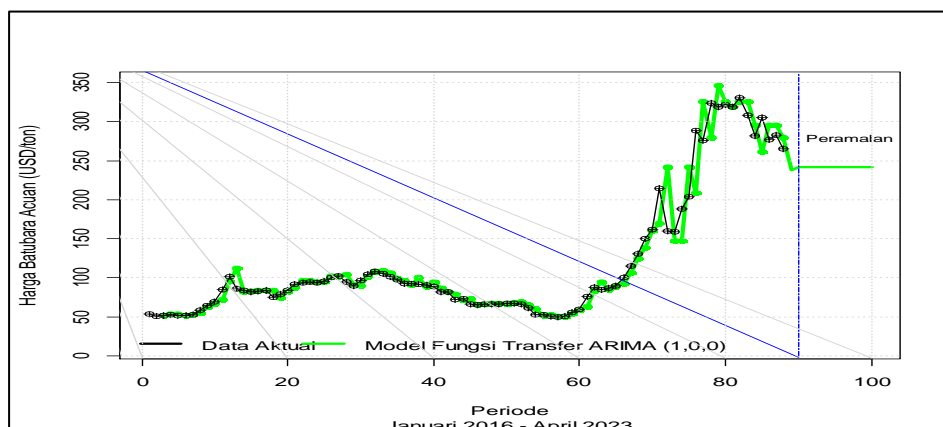


Figure 9.b. Transfer Function ARIMA (1,0,0)

$$\hat{Y}_t = (0,4198Y_{t-1} - 0,00001X_{3t-1} + 0,000004198X_{3t-2})$$

CONCLUSION

The transfer function model to forecast reference coal price are ARIMA (0,0,1) and ARIMA(1,0,0) are very accurate and forecasting for reference coal price tends stable. The results of the study obtained that the reference coal price in Indonesia is influenced by coal exports. The two transfer function models produce very good forecasts rate with MAPE 6.4069% and 6.2624%. The predicted results are close to the actual data pattern. Forecasting with the transfer function model shows that reference coal price in Indonesia for the period May 2023 to April 2024 tends to be stable.

ACKNOWLEDGEMENT

We acknowledgement to Faculty of Mathematics and Natural Sciences, Universitas Mulawarman which has supported this research.

REFERENCES

- [1] Akbar, J. *Ikan Papuyu Teknologi Manajemen dan Budi Daya*. (Universitas Lambung Mangkurat Press, Banjarmasin, Kalimantan Selatan, 2018).
- [2] J.E. Hanke and D.W. Wichern, *Business Forecasting Eight Edition*, New Jersey: Prentice-Hall, (2005).
- [3] R.J. Hyndman, and G. Athanasopoulos, *Forecasting: Principles and Practice*, Melbourne: OTexts, (2018).
- [4] S., Makridakis, S.C., Wheelwright, & V.E., Mc Gee, *Metode dan Aplikasi Peramalan Jilid I*, Edisi Kedua, Jakarta: Penerbit Erlangga, 1999.
- [5] D.C., Montgomery, C. L. Jenning, & M. Kulahci, *Introduction to Time Series Analysis and Forecasting*, New Jersey: Wiley-Interscience, (2008).
- [6] A., Umiyati, D., Dasari, & F., Agustina, Peramalan Harga Batubara Acuan Menggunakan Metode PSOSVR dan IPSOSVR. *Jurnal Sains dan Seni*, 9(1), (2020), 175–198.
- [7] W.W.S., Wei, *Time Series Analysis Univariate and Multivariate Methods* (second edition), Boston: Pearson Education Inc, (2006)
- [8] T. Anindita & A.A. Syaputra, Analisis Pengaruh Kurs USD, Harga Batu Bara Acuan, dan Volume Produksi Terhadap Volume Ekspor pada PT Bukit Asam (Persero)TBK, *Jurnal Manajemen Industri dan Logistik*, 1(2), (2017), 112-120.

Spatial Structural Equation Modeling of Poverty Indicators in the South Sumatra Region

Dyah Yunita Hartanti⁽¹⁾, Dian Agustina*⁽¹⁾, Winalia Agwil⁽¹⁾, Sigit Nugroho⁽¹⁾

⁽¹⁾Statistics Study Program, Bengkulu University, Indonesia

*Email: dianagustina@unib.ac.id

Abstract. Poverty is a condition where there is an inability to meet basic needs such as food, clothing, shelter, education and health. In the Southern Sumatra region, there are 3 provinces that have a poverty rate above 10%, namely South Sumatra Provinces, Bengkulu Provinces, and Lampung Provinces. Poverty in a region influenced simultaneously by various indicators from several variables, such as the economy, education, employment and others. Structural Equation Modeling (SEM) with the Partial Least Square approach is able to analyze patterns of relationships between latent constructs and their indicators, latent constructs with each other, as well as direct measurement errors. Spatial effect testing which shows that there is a relationship between regions directs us to the use of spatial SEM PLS. Then, testing the spatial effects which show that there is a relationship between regions, leads to the use spatial SEM PLS. Base on the results of the analysis, it is known that the spatial SEM model that used for modeling poverty in Southern Sumatra Region is the spatial lag model and the spatial error model with a queen contiguity matrix. In the spatial lag model, poverty is significantly affected by lag poverty and health. Meanwhile, in the spatial error model, poverty is significantly affected by the lag of poverty and health.

Keywords: Poverty, SEM-PLS, Spatial

INTRODUCTION

Poverty is a condition not only living in a lack of money and low income levels, but also many other things such as low levels of health, low education, unfair treatment in law, threats of crime, and powerlessness in determining their own way of life [2-1]. In addition, poverty is also caused by the large number of people who have limited access to basic services such as limited access to capital, production facilities, marketing, increasing the quantity and quality of products, sanitation, external influences such as fuel price fluctuations, tariffs and other regulations that cause an increase in the price of goods. and services as well as the increasingly limited ability of the population to meet their basic needs. From an economic perspective, the causes of poverty are low income, limited employment opportunities, slow economic growth and low levels of education.

Several provinces on Sumatra Island in 2021 have high poverty rates. In the South Sumatra region, there are 3 provinces that have a poverty percentage above 10%, namely South Sumatra with a poverty percentage of 12.84%, Bengkulu with a poverty percentage of 15.30%, and Lampung with a poverty percentage of 12.62% [3-2].

One of the methods that can be applied to model poverty in the South Sumatra region is Structural Equation Modeling (SEM) with the Partial Least Square approach. Because it is suspected that there is a spatial effect on the problem of poverty, the SEM-PLS spatial will be used. It is hoped that by providing a spatial effect on structural modeling it can provide validity to the relationship between each variable in the study so that the research results can provide an overview of the priority dimensions that must be addressed in poverty alleviation.

MATERIALS AND METHODS

Spatial SEM-PLS

Structural Equation Modeling with Partial Least Square (PLS) is a variant-based SEM method. If the sample data is limited in number, the data is not normally distributed multivariate, then the alternative used is variant-based SEM or also called Partial Least Square (PLS) [1-3]. The aim of SEM-PLS is to find predictive linear relationships between variables. The latent variable is defined as the sum of the composite weights of the indicators. The following is a form of structural model in the SEM-PLS method [4]

$$\eta = \mathbf{B}\eta + \mathbf{\Gamma}\xi + \zeta$$

with the measurement model as follows:

$$y = \Lambda_y \eta + \varepsilon$$

$$x = \Lambda_x \xi + \delta$$

Meanwhile, the spatial lag SEM model can be written as follows [5]:

$$\eta = \rho \eta_w + \Gamma \xi_x + \zeta$$

and the equation for the error spatial SEM is as follows:

$$\eta = \tau \eta_w + \Gamma \xi - \tau \Gamma \xi_w + \zeta$$

Research Method

The data in this study is secondary data obtained from the website of the BPS for the Southern Sumatra region which includes Jambi Provinces, South Sumatra Provinces, Bengkulu Provinces, Lampung and Bangka Belitung Islands. The data used is data on indicators as in the table below, covering the sectors of Poverty (Y_1), Economy (Y_2), Health (X_1) and Employment (X_2), with the observations being districts/cities in the Southern Sumatra region in 2021.

Table 1. The indicators of each latent variable

Latent Variable	Indicator
Poverty	Y_{11} District/City Poverty Levels in the South Sumatra Region
	Y_{12} Poverty Depth Index
	Y_{13} Poverty Severity Index
Economy	Y_{21} Percentage of households whose lighting source does not use electricity
	Y_{22} Percentage of households that use wood for cooking as fuel
	Y_{23} Percentage of households whose building ownership status is not their own
	Y_{24} Percentage of households whose building area is less than 19 m ²
	Y_{25} The percentage of households with the widest type of floor is land
Health	X_{11} Adequate Drinking Water
	X_{12} Percentage of population who have health insurance in the form of BPJS or Contribution Assistance Recipients
	X_{13} Percentage of households whose source of drinking water comes from unprotected wells/springs
	X_{14} Percentage of households that do not have public toilet facilities
Employment	X_{21} Number of Unemployed Individuals by Age Group with the Lowest 40% Welfare Status in the South Sumatra Region
	X_{22} Labor Force Participation Rate

Data analysis used the R application program. The following are the steps of the analysis carried out in this study.

1. Develop a research path diagram.
2. Designing measurement models and structural models.
3. Evaluate the suitability of the model (Goodness of Fit).
4. Calculates factor scores for each latent variable.
5. Form a spatial weight matrix (Queen Contiguity and Rook Contiguity)
6. Perform spatial autocorrelation tests with Moran's I statistics on latent variable scores obtained from estimating structural models in SEM.
7. Select the best spatial structural model by looking at the results of the Lagrange Multiplier test.
8. Interpret and conclude results.

RESULT AND DISCUSSION

The conceptual framework built in this study is described in the following figure.

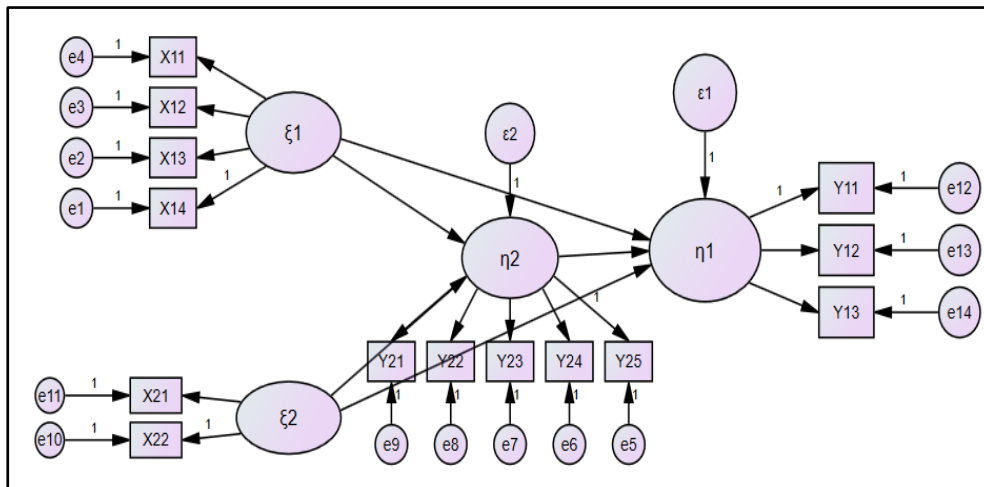


Figure 1. The Path Diagram for the model

Validity Test

The individual reflective measure is said to be valid if the loading factor value is ≥ 0.4 for each indicator of the latent variable. Based on the initial validity test it was found that $(X_{2,2})$, $(X_{1,2})$, $(X_{1,3})$, $(Y_{2,1})$, $(Y_{2,4})$ had a loading factor value < 0.4 which was then excluded from the model. The results of the discriminant validity test of the measurement model can be seen in Table 2. Based on these results it is known that the root value of the average variance extracted (Average Variance Extracted) of all latent variables is greater than the correlation value between latent variables. So it can be concluded that the indicators used in this study have fulfilled the discriminant validity test.

Table 2. Discriminant validity test results

Latent Variable	Correlation				AVE root
	Employment	Health	Economy	Poverty	
Employment	1,000	0,221	0,358	0,216	1,000
Health	0,221	1,000	0,188	0,193	0,720
Economy	0,358	0,188	1,000	0,359	0,849
Poverty	0, 216	0,193	0,359	1,000	0,942

Reliability Test

Composite reliability is used to test the reliability of latent variables. A composite reliability value of more than 0.7 indicates that the latent variable has a good reliability value, but a value of 0.6 is still acceptable. The reliability test results show that all composite reliability values are in the indicator blocks that make up the latent variables Health, Economy, Employment and Poverty has a fairly good reliability. Then an evaluation of the structural model was carried out by looking at the Goodness of Fit value obtained at 0.3272, meaning that the exogenous latent variables consisting of Employment and Health were able to explain 32.72% of the diversity in Economy and Poverty and the rest was explained by other variables.

Table 3. Estimation results of structural model parameters

	Coefficient	SE	p-value	t-stat
Employment -> Poverty	0,0789	0,126	0,09*	1,681
Employment -> Economy	0,3329	0,121	0,01**	2,76
Health -> Economy	0,1143	0,123	0,00**	-11,11
Health -> Poverty	0,1180	0,178	0,04**	-2,095
Economy -> Poverty	0,3088	0,136	0,01**	2,698

The following model is a structural model that is formed:

$$\text{Economy} = 0,1143\text{Health} + 0,3329\text{Employment}$$

$$\text{Poverty} = 0,3088\text{Economy} + 0,1180\text{Health} + 0,0789\text{Employment}$$

Spatial Effect Test

Moran's I test is used to see whether there are spatial dependencies between locations or not. The hypothesis for testing Moran's I is

$$H_0 : I = 0 \text{ (tidak ada autokorelasi spasial antar lokasi)}$$

$$H_1 : I \neq 0 \text{ (ada autokorelasi spasial antar lokasi)}$$

Based on the results of Moran's I test using the queen contiguity matrix, it shows that at a significance level of 5%, the value $Z(I) = 9.39 > Z_{\text{table}} = 1.96$. This means that H_0 is rejected, meaning that there is spatial autocorrelation between locations. While the results of the Moran's I test using the rook continuity matrix also show that at a significance level of 5%, the value $Z(I) = 8.27 > Z_{\text{table}} = 1.96$. This means that H_0 is rejected, meaning that there is spatial autocorrelation between locations. So it can be concluded that by using the two matrices, there is an autocorrelation or spatial dependence on poverty in the Southern Sumatra Region.

Lagrange Multiplier Test

Table 4. Spatial effect test for the poverty model

Spatial Dependencies Test	Queen Contiguity Matrices		Rook Contiguity Matrices	
	Statistics	p-value	Statistics	p-value
Lagrange Multiplier (lag)	81,6	0,000**	68,54	0,000**
Lagrange Multiplier (error)	68,75	0,000**	51,65	0,000**

a. Lagrange Multiplier Test Results with Queen Contiguity Matrices

The Lagrange Multiplier (lag) test aims to identify inter-regional linkages. Based on the processing results in Table 4, the p-value for LM-lag is 2.2×10^{-16} or less than $\alpha=0.05$. So it can be concluded that reject H_0 , meaning that there is a spatial lag dependency. The Lagrange Multiplier (error) test aims to identify the relationship between errors between regions, based on the processing results in Table 4, the p-value for the LM-error is 2.2×10^{-16} or less than $\alpha=0, 05$. So it can be concluded that reject H_0 , meaning that there is a spatial dependency error.

b. Lagrange Multiplier Test Results with Rook Contiguity Matrix

The Lagrange Multiplier (lag) test aims to identify links between regions. Based on the processing results in Table 4, the p-value for LM-lag is 2.2×10^{-16} or less than $\alpha=0.05$. So it can be concluded that reject H_0 , meaning that there is a spatial lag dependency. The Lagrange Multiplier (error) test aims to identify the relationship between errors between regions, based on the processing results in Table 4, the p-value for the LM-error is 6.64×10^{-13} or less than $\alpha=0, 05$. So it can be concluded that reject H_0 , meaning that there is a spatial dependency error.

Best Model Selection

The selection of the best model in this study uses the Akaike Information Criterion (AIC) value. The model that has the smallest AIC value is the best model that can be used in modeling poverty in the Southern Sumatra Region. From Table 5, it is known that the SAR model and SEM model using a queen contiguity weighting matrix have a smaller AIC value than the rook continuity weighting matrix so that the structural model chosen is the SAR and SEM model with a queen contiguity weighting matrices.

Table 5. AIC criteria on the model

Latent Variable	AIC value			
	SAR Model		SEM Model	
	Queen Contiguity	Rook Contiguity	Queen Contiguity	Rook Contiguity
Poverty	39,493	62,204	43,681	62,577

Spatial Lag and Spatial Error Models in SEM

Table 6. Estimation of SAR structural model parameters

Latent Variable	Queen Contiguity Matrices			
	SAR Model		SEM Model	
	Coefficient	p-value	Coefficient	p-value
Economy	0,0967	0,21	0,1262	0,23
Health	0,1021	0,1*	0,2264	0,02**
Employment	0,0071	0,91	0,0752	0,51
Lag Poverty	0,8722	0,00**	0,8782	0,00**
Lag Health			0,1112	0,22
Lag Economy			0,2957	0,19
Lag Employment			-1,1873	0,24

The spatial lag model in the SEM for poverty in the Southern Sumatra region can be written as:

$$Poverty_i = 0,8722 \sum_{j=1, i \neq j}^n w_{ij} Poverty_j + 0,0967 Economy_i + 0,1021 Health + 0,0071 Employment_i$$

In the model above, it is known that the rho value is 0.8722 and is significant at the 5% level of significance, this means that there is a correlation between the poverty rate in one region and other adjacent areas. Meanwhile, the spatial error model in the SEM for poverty in the Southern Sumatra Region can be written as follows:

$$Poverty_i = 0,8782 \sum_{j=1, i \neq j}^n w_{ij} Poverty_j + 0,1262 Economy_i + 0,2264 Health_i + 0,0752 Employment_i - 0,2957 \sum_{j=1, i \neq j}^n w_{ij} Economy_i - 0,1111 \sum_{j=1, i \neq j}^n w_{ij} Health_i + 0,1873 \sum_{j=1, i \neq j}^n w_{ij} Employment_i$$

In the model above, it is known that the lambda value is 0.8782 and is significant at the 5% significance level, this means that there is a link between the poverty rate in one region and other adjacent areas.

CONCLUSION

Based on the results and discussion, it can be concluded that the SEM spatial model that is significant for poverty in the South Sumatra region based on the lagrange multiplier test is spatial autoregressive and spatial error using a spatial queen contiguity weighting matrix. In the spatial autoregressive SEM model, the variables that affect poverty in a city/district are poverty in neighboring cities/regencies, economy, and health. Meanwhile, in the SEM spatial error model, the variables that affect poverty in a city/district area are poverty in neighboring cities/regencies, health, economic conditions in neighboring cities/regencies, health in neighboring cities/regencies, and employment in neighboring cities/regencies.

ACKNOWLEDGMENT

Thank you to the University of Bengkulu for supporting this research through the Fundamental Scheme of the Institute for Research and Community Service.

REFERENCES

- [1] Suryawati, C. 2005. Memahami Kemiskinan Secara Multidimensional. *JMPK*, Vol. 08 no 1. pp. 121-129.
- [2] Badan Pusat Statistik. 2021. *Indikator Kemiskinan*. BPS: Indonesia.

-
- [3] Ghozali, I. 2008. *Structural Equation Modelling Metode Alternatif dengan Partial Least Square (PLS)*. Semarang: Badan Penerbit Universitas Diponegoro.
 - [4] Chin, W. W. (1998). *The Partial Least Squares Approach to Structural Equation Modeling*. Dalam G. A. Marcoulides, *Modern Methods for Business Research* (hal. 295-336). London: Lawrence Erlbaum Associates
 - [5] Oud, J., dan Folmer, H. 2008. *A Structural Equation Approach to Spatial Dependence Models*. Behavioural Science Institute, Radbound University Nijmegen.

PROCEEDING

THE 6TH INTERNATIONAL CONFERENCE ON SCIENCE AND TECHNOLOGY (ICST)



ISSN 3031-5514



9 773031 551001